

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 12:52:35 ; Search time 514 Seconds
(without alignments)
10206.037 Million cell updates/sec

Title: US-10-047-855-4
Perfect score: 3206
Sequence: 1 gtcgaccacgcgtccggc.....aaaaaaaaaaggcgccgc 3206

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	68.4	2.1	1180	4	US-09-270-767-27552
C 2	68.4	2.1	2786	4	US-09-270-767-11893
C 3	59.6	1.9	665	2	US-08-883-795A-36
C 4	57.8	1.8	187169	4	US-09-949-016-12776
C 5	57.8	1.8	191569	4	US-09-949-016-15940
C 6	57.2	1.8	1141	4	US-09-806-708B-22
C 7	54.8	1.7	3042	3	US-09-255-829-17
C 8	54	1.7	832	4	US-09-621-976-2813
C 9	53	1.7	88490	4	US-09-949-016-12758
C 10	53	1.7	88736	4	US-09-949-016-14222
C 11	52	1.6	3404	4	US-09-710-279-4346
C 12	51.6	1.6	1987	4	US-09-350-756-8
C 13	51.6	1.6	2613	3	US-09-255-829-7
C 14	51.6	1.6	2616	3	US-09-255-829-1
C 15	51.6	1.6	2622	3	US-09-255-829-25
C 16	51.6	1.6	2622	3	US-09-255-829-5
C 17	51.6	1.6	2628	3	US-09-255-829-9
C 18	51.6	1.6	2637	3	US-09-255-829-11
C 19	51.6	1.6	2685	3	US-09-255-829-3
C 20	51.6	1.6	2724	3	US-09-255-829-15
C 21	51.6	1.6	2862	3	US-09-255-829-13
C 22	51.6	1.6	3891	1	US-08-480-604A-27
C 23	51.6	1.6	3891	2	US-08-405-496A-27
C 24	51.6	1.6	3891	3	US-08-915-136-27
C 25	51.6	1.6	3891	4	US-09-084-517-27
C 26	51.6	1.6	4835	4	US-09-288-326A-10
C 27	51.4	1.6	12313	4	US-09-949-016-13248

BEST AVAILABLE COPY

C	28	51	1.6	1141	4	US-09-806-708B-22	Sequence 22, Appl
	29	50.4	1.6	2421	3	US-09-248-796A-1858	Sequence 1858, Ap
	30	50.4	1.6	6124	3	US-08-213-419B-3	Sequence 3, Appl
	31	49.4	1.5	119153	4	US-09-949-016-12378	Sequence 12378, A
	32	49.2	1.5	108310	4	US-09-949-016-16366	Sequence 16366, A
	33	49	1.5	601	4	US-09-949-016-156535	Sequence 156535,
	34	48.6	1.5	29717	4	US-09-949-016-16284	Sequence 16284, A
C	35	48.6	1.5	60376	4	US-09-949-016-12423	Sequence 12423, A
	36	48.2	1.5	251672	4	US-09-949-016-17296	Sequence 17296, A
	37	48.2	1.5	251682	4	US-09-949-016-11973	Sequence 11973, A
C	38	47.8	1.5	231129	4	US-09-949-016-16110	Sequence 16110, A
	39	47.8	1.5	266293	4	US-09-949-016-11934	Sequence 11934, A
C	40	47.4	1.5	9447	4	US-09-949-016-14649	Sequence 14649, A
	41	47.4	1.5	50383	4	US-09-949-016-17600	Sequence 17600, A
C	42	47.2	1.5	601	4	US-09-949-016-54020	Sequence 54020, A
C	43	47.2	1.5	601	4	US-09-949-016-54021	Sequence 54021, A
C	44	47.2	1.5	601	4	US-09-949-016-54024	Sequence 54024, A
C	45	47.2	1.5	601	4	US-09-949-016-108654	Sequence 108654,

ALIGNMENTS

RESULT 1

US-09-270-767-27552/c
; Sequence 27552, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27552
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-27552

Query Match	2.1%	Score 68.4;	DB 4;	Length 1180;
Best Local Similarity	44.7%	Pred. No. 1.2e-07;		
Matches	409;	Conservative	0;	Mismatches 491;
				Indels 16;
				Gaps 3;
QY	978	CAGAAATTCGCGAAACAATAGCAAGTACAGAGTACACTATATAATATTAAAGCCATT	1037	
DB	920	CAAGATGCAGATGAGACCCCTGGCCAGGCTGAGGCTTCCCTATGTGCGGTGCGACCTTA	861	
QY	1038	ACCAGGATACAGTTGTGACATGAAATCTTCAATTTTCCAAAGTATTGGAGCCAAAGATACC	1097	
DB	860	CQGCTACTCGCGCGTAGACTTCAAGAACACCTATGTCTCACTACTTGGCCCAAGAGCTGGCC	801	
QY	1098	ATTGATGTTGGCCATCGAGTGCAGGAACTCTACAACTCCAGCTGGCTAAAGT	1157	
DB	800	CAACCTGGATGGGTCTATCGTGGAAATGGCAAGATTACATTGCACGCTCTCTCGCGA	741	
QY	1158	TCAAGAAATACTATTGCTCTTTAAGAAATGCTAGTATCATGTCAGCCTTTGTAGA	1217	
DB	740	AAGGAGAACACTATCGGGCTTTCTTGAGGCCCATGAACATCAGCAGACATGATCGA	681	
QY	1218	ATTGACGTACACCTTTTCAAGGACTTTGTGCGCGTGGTATATCATGATCTTACCTGTG	1277	
DB	680	GTTGGATGCTCATTTGACTGCTGATGGTGTGCTGCTGTGATTTATCAGATTTTGGGACTCG	621	
QY	1278	TTTGACTATGAAAAGAAATTTGATGCTGATCCAGTGAATTTTGAATTTCCAGTAAA	1337	
DB	620	AACTGCTCCGCTCGCAAGCAGATCAGCCGCGAGACAGCTGGAGTAGTCGCTGATCAA	561	
QY	1338	AGAATTAACATTTCACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTGAAATC	1397	
DB	560	AGACATAACTATGAGCTGCTCAAAAGCTCGGCATCTTCTGTGATCGCAGGCCAAGT	501	

Qy	1398	TAAGATCGGAAGAATCTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAATCAGCCATT	1457
Db	500	GAGGAG-----TATCCTCGCACACAGCCGAGCCAGGATGGAACACCGCATATT	450
Qy	1458	TCCTTCTCTTA-AGATGGTTTTCAGAGTCTTTGCCAGAAAGATGTAGGGTTTAAACATTGAAA	1516
Db	449	CCCAACGCTGGTTGGAGTACTGAGAAAGCTGCCCAAGTCACTGGGCATTGATGTGGAGA	390
Qy	1517	TAAATGGATCGCCAGCAAGGATGGAATGTGGGATGGTAATTAACATATATTTTG	1576
Db	389	TTAAGTGGCCACAGCGTCGCCAGGCGGAGGATCAGAGGCTGAGCAACAA-----TCG	336
Qy	1577	ACATGAATCTGTTTGGATATATATTTTAAAACTGTGTTTAAAGAAATTCGGGAAGAGGA	1636
Db	335	ACAAGAACTTCTCGCCACAGGTGATCCATCAGGTGATCCAGAGGGCTGTGGCAGAC	276
Qy	1637	GAAATGCTTTTCTTTCATTCATGACAGATATTTGCACAAATGGTTTCGGCAAAAGCAGAAACA	1696
Db	275	CGATAATCTTCTCAGCTTCATGCTGACATGTGCAGATGCTCAGGTTCAGAGCAGAAACG	216
Qy	1697	AATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTATCTTGAATCTCATGGACC	1756
Db	215	TCCTCCAGTGTGTTCTCAGCAGGAGAGAGCAAGAAAGTGGCAGCGCTTCTCTGGATC	156
Qy	1757	TCAGATCTCGGACAAACCCCATTCGAATGAGCTTTGCACAGTTTGAAAATCTACTGGGGA	1816
Db	155	TGCGAACACGAGACCTTTATTGACGCGGTAAACAAATGCTCAAGCTTCGAGCTGGCTGGCA	96
Qy	1817	TAAATGTACATCTCAAGACTTGTCTCAGAAACCCATCTATATTCAGAGGCAAAAGCTA	1876
Db	95	CAGCTCCACAGCCGAGGACTTCTTGGGCGAAATGCAATCAGAAATGCTGCGAAAGCCA	36
Qy	1877	AGGGACTAGTCATATT 1892	
Db	35	AGGATCTGGCCAAAT 20	

RESULT 2
US-09-270-767-11893/c
; Sequence 11893, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11893
; LENGTH: 2786
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11893

Qy	978	CAGAAATCCCGAAACAAATAGGCAAGTGTGAGAGTTGACTATATATTAATTAAGCCATT	1037
Db	1895	CAAGAAATGCAGATGAGACCCCTGGCCAGGCTGAGGCTTCCCTATGTGGCGGTGACGCCCTTA	1836
Qy	1038	ACAGGATACAGTTGTGACATGAATCTTCATTTCCAAAGTATTTGAAGCCCAAGATACC	1097
Db	1835	CCGCTACTCGCCGCTAGACTTCAGAAACACCTAATGCTACTGCTGCCCCAAGGCTGGCC	1776
Qy	1098	ATTGGATGTGGCATCGAGGTGAGGAACTCTCAACAACTGCCAGCTGGCTAAAGT	1157
Db	1775	CAACCTGGATGTGGTCATCTGTGAAATGCGAAGATTACATTGCAGACGCTCTCGCGGA	1716
Qy	1158	TCAAGAAATPACTATTGCTTCTTTAAGAAATGCTGCTAGTCATGTTGCGAGCCTTTGTAGA	1217

Db	1715	AAGGAGAACACTATTCGGTCTTTCTCTGAGCGCCCATGAACATCAGCAGACATGATCGA	1656
Qy	1218	ATTTCACGTACACCTTTCAAAGGACTTTTGTCCCGTGGTATATCATGATCTTACTGTTG	1277
Db	1655	GTTGATGTCAATTGACTGTGATGGTGTGCTGTGATTTATCAGGATTTCCGACTGCG	1596
Qy	1278	TTTGATATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTTGAATTCAGTAAA	1337
Db	1595	AACTGCTCGCCTGGCAAGCAGATCAGCCGCCAGACCCAGCTGGAGTACGTGCTGATCAA	1536
Qy	1338	AGAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTGAAATC	1397
Db	1535	AGACATAAATATGAGCTGCTCAAAAGGCTGGCATCTTCTCTGTGATCGCAGGCCAAGT	1476
Qy	1398	TAAAGTCGAAAGAATCTGTGTTTCAGAGGAAAAATTCCTTTTCAGAAAAATCAGCCATT	1457
Db	1475	GAGGAG-----TATCCCTCCCAACACGCCAGCCAGGATGGAACACCCGATATT	1425
Qy	1458	TCCTTCTCTTA-AGATGGTTTTCAGAGTCTTTGCCAGAAAGATGTAGGGTTTAAACATTGAAA	1516
Db	1424	CCCAACGCTGGTTGGAGTACTGGAGAAAGTCTGCCCAAGTCACTGGGCATTTGATGTGGAGA	1365
Qy	1517	TAAATGGATCTGCCAGCAAGGATGGAATGTGGGATGGTAACCTTATCAACATATTTTG	1576
Db	1364	TTAAGTGGCCACAGCGTTCGCCAGGCGGAGGATCAGAGGCTGAGCAACAA-----TCG	1311
Qy	1577	ACATGAATCTGTTTTCAGATATAATTTTAAAACTGTTTTCAGAAAAATTCGGGAAGAGA	1636
Db	1310	ACAAGAACTTCTTCCGCGACAAAGGTATCCATCAGGTGATCCAGAAAGGCTGTGGCAGAC	1251
Qy	1637	GAAATGTTTCTTTCATTTGATGACAGATATTTGCAATGTTGCAATGTTCCGCAAAAGCAGAAACA	1696
Db	1250	CGATAATCTTCCAGCTTCGATGCTGACATGTGCAGGATGCTGAGGTTCAAGCAGAAACG	1191
Qy	1697	AATATCGATATATTTTAACTCAAGGAAATCTGAGATTTATCTGAACTCATGGACC	1756
Db	1190	TCCTCCAGTGTGTTCTCAGCAGGAGAGACGAAAGATGGCAGCGCTTCTCTGGATC	1131
Qy	1757	TCAGATCTGGCAACACCCCATTCGCAATGAGCTTTGCACAGTTTGAATAATCTACTGGGGA	1816
Db	1130	TGCGAACACGAGCCTTTTATTCAGCCGTTAAACATGCTCAAGCTTTCGAGCTGGCTGGCA	1071
Qy	1817	TAAATGTACATCTCAAGACTTGTCTCAGAAACCCATCTTATATTCAGAGGCAAAAGCTA	1876
Db	1070	CAGCTCCACAGCCGAGGACTTCTTCCGCGAAATGCAATGCTGCGAAATGCTGCGAAAGCCA	1011
Qy	1877	AGGACTAGTCATATT 1892	
Db	1010	AGGATCTGGGCCAAAT 995	

RESULT 3
US-08-883-795A-36
; Sequence 36, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awarig, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: RH 32
US-08-883-795A-36

Query Match 1.9%; Score 59.6; DB 2; Length 665;
Best Local Similarity 47.1%; Pred. No. 1.9e-05;
Matches 214; Conservative 0; Mismatches 239; Indels 1; Gaps 1;

```
QY 2212 TATTCATTTTTCATCACTGAGCAATGCTGATCTATGCGCTTTGGGCTTCTCAGTTCAATG 2271
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66 TATTTATAATTAAATATTTTATAATTAAATATTTATAATTAAATATTTATAATTAAAT 125

QY 2272 AAGCAATAATGAAGTATTTAACTCTTTCACTA-CAGTCTTCGCAAGTATGCTATTTAAAT 2330
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 126 TATTTATAATTAAATATTTTATAATTAAATATTTATAATTAAATATTTATAATTAAAT 185

QY 2331 TACTTGGCCAGGTAAATTTGCGAGTCGCTCTTTATAGTGAGAAAATTTATGGTGTAGT 2390
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 186 TATTTATAATTAAATATTTTATAATTAAATATTTATAATTAAATATTTATAATTAAAT 245

QY 2391 AATATAAATATTTTAACTAAATATATAAATCTATAATGCTTAAACATATGTTCAATAAA 2450
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 246 TATTTATAATTAAATATTTTATAATTAAATGTTTATAATTAAATATTTATAATTAAAT 305

QY 2451 GCATGACCTTTGAAATTAACATATATAAATAGCTCATATTTTACCTTACAGCTTTTCAT 2510
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 306 TGTTTATAATTACATATTTTATAATTAAATGTTTATAATTACATATTTTATAATTAAAT 365

QY 2511 TGATCAGGTCTGAATCTTTAGCACTTAAGGAAAATGACTATGCATAATTATACCTGACC 2570
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 366 TGTTTATAATTACATATTTTATAATTAAATGTTTATAATTACATATTTTATAATTAAAT 425

QY 2571 ATGAAAAAATAAGTACCTCAATGCATGCAATTCGACCTGGTATTCGACCTGCACAAAT 2630
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 426 TGTTTATAATTACATATTTTATAATTACATATTTTATAAAGTATTTATAATTACATATTT 485

QY 2631 CTTTGTGCCATCTTGTATATAGGTATTTTATACA 2664
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 486 TATAATTAAAGTATTTATAATTACATATTTTATA 519
```

RESULT 4

US-09-949-016-12776
Sequence 12776, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12776
LENGTH: 187169
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(187169)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12776

Query Match 1.8%; Score 57.8; DB 4; Length 187169;
Best Local Similarity 43.6%; Pred. No. 0.0012;
Matches 257; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

```
QY 2361 TCTTTATAGTCAGAAAATTTATTGGTTAGTAATAATAATTTTAAACTAATAATATAAA 2420
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 624 TATATAATATAATAAATATATATAATATAATAATAATAATAATAATAATAATAATAT 683

QY 2421 TCTATAATGTTTAAACATATGTTTCATTTAAAGCATAGCACCTTTGAAATTAACATATAATA 2480
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 684 ACATAAATAATATAATAATATAATAATAATAATAATAATAATAATAATAATAATAATA 743

QY 2481 AGCTCATATTTACACTTACAGCTTTTCATTTTGATCAGGCTCGAAATCTTTAGCAGCTTAAG 2540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 744 AAAATATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 803

QY 2541 GAAATGACTATGCTATATTAACCTGACCATGAAABAAAATAGCTCAATGCCATGC 2600
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 804 ATATATAAATAATAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 863

QY 2601 ATTTGCACTGCTGATTTCCAACTGCACAAATCTTTGTGCGCATCTTGTATATAGGTATTTT 2660
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 864 AATATAATAAATAATAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 923

QY 2661 TACATGGTTCACATGACACACACACACATTTTCATTCAGTATGAACTTTGAGGCTGCTGC 2720
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 924 TATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 983

QY 2721 CATTTTCCACTTAAACCAACCCAGCTGAAGTCAACCTCGAAACTTGTTCATAAATCT 2780
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 984 TGATATATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1043

QY 2781 TTCAAAAGTTGTTTACATCAATGTTAAAAATTTCAAAATGCTGCAGGTAATTTAAATGTA 2840
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1044 TATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1103

QY 2841 TAAATAATTAGTAAGAAAAAGTATGTATTCATCTTAGTATAGATAGATAGATAGATAGAT 2900
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1104 TATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1163

QY 2901 AATTCAATTCAGTGCATGCTTTAGTCTTAAGCATGAGATTGTACATGT 2949
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1164 TAATACATATAATATATATATATATATATATATATATATATATATATATATATATATAT 1212
```

RESULT 5

US-09-949-016-15940
Sequence 15940, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755									
; PRIOR FILING DATE: 2000-10-20									
; PRIOR APPLICATION NUMBER: 60/237,768									
; PRIOR FILING DATE: 2000-10-03									
; PRIOR APPLICATION NUMBER: 60/231,498									
; PRIOR FILING DATE: 2000-09-08									
; NUMBER OF SEQ ID NOS: 207012									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 15940									
; LENGTH: 191569									
; TYPE: DNA									
; ORGANISM: Human									
; FEATURE:									
; NAME/KEY: misc_feature									
; LOCATION: (1)...(191569)									
; OTHER INFORMATION: n = A,T,C or G									
US-09-949-016-15940									
Query Match									
Best Local Similarity 43.6%; Score 57.8; DB 4; Length 191569;									
Matches 257; Conservative 0; Mismatches 332; Indels 0; Gaps 0;									
QY	2361	TCTTTATAGTCGAGAAATTTATTGGTTAGTAAATATAATATTTTAAACTAAATATATAAA	2420						
DB	624	TATATAATATAATAATATAATAATATAATAATATAATAATATAATAATATAATATAAT	683						
QY	2421	TCTATAATGTTAAACATATGTTTCATTAAAGCATAGCACTTTTGAAATTAACATATAAAT	2480						
DB	684	ACATAAATATAATAATATAATATAATATAATAATAATAATAATAATAATAATAATAA	743						
QY	2481	AGCTCATATTTACACTACAGCTTTTCATTTGATCAGGTCTGAAATCTTTAGCACTTAAG	2540						
DB	744	AAAAATATAAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	803						
QY	2541	GAAATGACTATGCATATTAACCTGACCACTGAAAGAAATAAGTACCTCAAAATGCATGC	2600						
DB	804	ATATATAAATATAATAAATATAATAATAATAATAATAATAATAATAATAATAATAATAA	863						
QY	2601	ATTTCGACTGGTGATTCCCACTGCACAAATCTTTGTGGCATCTTGTATATATAGTATTTT	2660						
DB	864	AATATATAAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	923						
QY	2661	TACATGGGTGATGCACACAAACCAATTTTCATTCAGTATGAACTTGAGCGTGCCTGC	2720						
DB	924	TATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	983						
QY	2721	CATTTTCCACTTAACCAACAGCTGAGGTGACCTCGAAACTTGTTCATAAACTCT	2780						
DB	984	TGATATAATAAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	1043						
QY	2781	TTCAAAAGTTGTTTACATCAATGTTTAAATTTTCAAAATGCTGCAGGGTAATTTAATGTA	2840						
DB	1044	TATATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	1103						
QY	2841	TAAATATATTAGTAAGAAAAAGTATGATTTGCACTACTTAGTAGAATAGATCAACATACA	2900						
DB	1104	TATAAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	1163						
QY	2901	AATTCATTTAGTGCATGCTTTAGGTGTTAAGCATGAGATTTGACATGT	2949						
DB	1164	TAATACAT	1212						
RESULT 6									
US-09-806-708B-22									
; Sequence 22, Application US/09806708B									
; Patent No. 6784342									
; GENERAL INFORMATION:									
; APPLICANT: The University of British Columbia									
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants									
; FILE REFERENCE: 4810-58741									
; CURRENT APPLICATION NUMBER: US/09/806, 708B									
; CURRENT FILING DATE: 2001-04-03									
; PRIOR APPLICATION NUMBER: US 60/147,133									
; PRIOR FILING DATE: 1999-08-04									
; NUMBER OF SEQ ID NOS: 23									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 22									
; LENGTH: 1141									
; TYPE: DNA									
; ORGANISM: Artificial sequence									
; FEATURE:									
; NAME/KEY: promoter									
; LOCATION: (1)..(1141)									
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEl promoters									
US-09-806-708B-22									
Query Match									
Best Local Similarity 11.2%; Score 57.2; DB 4; Length 1141;									
Matches 121; Conservative 395; Mismatches 556; Indels 11; Gaps 3;									
QY	1841	TCAGAAACCCATCTATATTCAAGAGGCAAGCAAGCTTAAGGACCTAGTACATATCTGCTGGG	1900						
DB	60	THARMYCKYRRWYNNKSRWKGWYKKWYBCANNTSBRVHARRWKDKMTAYBMTWTKWG	119						
QY	1901	GTGATGATACCAATGATCTCTGAAAAAGAGAAATTTGAAGAACTTGGAGTTAATGGTC	1960						
DB	120	KTCGRHRYWRWAMBDTVDDHHYVTAMNNATWTCMMDKDDKTRWKKNNATGDDDT	179						
QY	1961	TAATTTATGATAGATATATGATGATGCTGCAACCAACCAATATATTCCAAGTGGAGC	2020						
DB	180	KYHMNNNGCBVTVMVRYKTDSDSKRMNYGMBWKNWSYDVVTYVWVDDMCKRKYR	239						
QY	2021	AATTTGAACCCCTGAAGAGGAAATTCGACAGCTTAAGAGCTGTTGTGTCCACCTGTTA	2080						
DB	240	RWRTGRMKNYVWVWETAHRRYNNGTWTEAMAYRWTWNNNNNNNAKAKRKYWGNR	299						
QY	2081	GCCTGTTTGTCCCTCATCTTTGTGTGGGAGTCTGATATATCCATG--TGGATGCCAACGG	2138						
DB	300	ABVNSTCTTWKSKTTKVRTSWANNCRAGDANKDHKWKWSAAMGVYNNNNNNNTYK	359						
QY	2139	CATTGATAACCTGGAGATGCTTAGTTTTATTGACAGAGGTCATTTTGGGGCGCTGCA	2198						
DB	360	ARHBARDMWMSAKKWHANAHAHYSRKKWTBTKRMTVNNNNNGTTWKKMWAYWKMMD	419						
QY	2199	CCGCTGTTCTGGGTATTCACTTTTTCATCACTGAGCATTTGTGATCTATGCTTTTGGCT	2258						
DB	420	DWBTYNNNNNGRTYYGWTNKKWYTKKANNCKWRADWDHKTCTHNNTTWKKMTY	479						
QY	2259	TCTCATGTTCAATGAAGCAATAATGAAGTATTTAACTCTTTTCACTACAGTTCTTGCAGTA	2318						
DB	480	NNCYKSWTNGKSHRBAAAVVYTWMMWRRYAHANNNNNDYMWKACTWYKYVCYCKWNN	539						
QY	2319	TGCTATTAAATTAATCTTGGCCAGGTATAATTCGACGTCAGTCTCTCTTTTATAGGAAAAA	2378						
DB	540	YAAWYTKSWNTYSRYRWKTNNSWRWSRSDTRSMGRANNYARABHY----GYKNWTRWB	595						
QY	2379	TTATTGGTTAGTATAATAATAATTTTAACTAAATATAATAATACTATAATGTTTAAACATA	2438						
DB	596	WSHTWBHBRAGAHHYMBMYBAKCHCKWYKAKKYAGAGSNNNNNNNNNNNNNNNNNA	655						
QY	2439	TGTTTCATTAAGACATAGCACCTTTGAAATTAACCTATATAAATAGCTCATATTATACCTTA	2498						
DB	656	TCARDDYYAASRWYMANAKWYKBAANNAYYTHANNWGMNNAATDTRRTWKKNNNN	715						
QY	2499	CAGCTTTTCATTGATCAGGTCTGAAATCTTTAGCACCTTAAGGAAATGACATATGATAA	2558						
DB	716	NAGTWKNNNNNNNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNA	775						
QY	2559	TTATACCTGACCATGAAABAAATAGTACCTCAAAATGCATGCTGCTTGTGACCTGGTATCC	2618						
DB	776	AYTKYTTNNNNYRGVVNTTAARDGWANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	835						
QY	2619	AACTGCACAAATCTTTGTGCACTCTTGTATATAGGATTTTTCATGAGGTTTGCATGCA	2678						
DB	836	NNNNNNNNNAYAWWTKWYTTDDWRBAYTTNNNNNNNNNMYYGAYADDYATYNSDTCDAW	895						


```
; LENGTH: 3404
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4346

Query Match
Best Local Similarity 1.6%; Score 52; DB 4; Length 3404;
Matches 139; Conservative 0; Mismatches 105; Indels 4; Gaps 2;

QY 2342 GTATATTTCCAGTCAGTCTCTTTATAGTCAGAAATTTATTTGGTAGTATATAATAT 2401
DB 416 GAAATAATTAAGATCAAAATCTTAATATGTTGTAATTTAATGGTACCTAAATAA 475
QY 2402 TTAAACTA-ATATATATAATCTATAATGTTAAACATATGTTCAATAAA--GCATAGC 2457
DB 476 TTAAACTACAATGCTAAATCCATATGTTGTTTCATTAGAGATTTAAATAATGATTATA 535
QY 2458 ACTTTGAAATTAACATATATAATAGCTCATATTTACACTTACAGCTTTTCATTTGATCAG 2517
DB 536 ACACATAAAGATTTCAAATTTATTTATTAATATAAATTTACATATGATAACGAATAAC 595
QY 2518 GTCTGAATCTTTAGCAGCTTAAGGAAATGACTATGCAATATATACCTGACCATGAAAA 2577
DB 596 AATTCCTAATAAATTTATTTTGTGATTTATTTTATTAATATATATATATATGAAAA 655
QY 2578 AATAAGT 2585
DB 656 ATATAAGT 663

RESULT 12
US-09-350-756-8
; Sequence 8, Application US/09350756
; Patent No. 6495143
; GENERAL INFORMATION:
; APPLICANT: U.S. Army Medical Research Institute for Infectious Diseases
; APPLICANT: John S. Lee
; APPLICANT: Peter Pushko
; APPLICANT: Michael D. Parker
; APPLICANT: Jonathan P. Smith
; APPLICANT: Mark T. Dertzbaugh
; APPLICANT: Leonard Smith
; TITLE OF INVENTION: Botulinum Neurotoxin Vaccine
; FILE REFERENCE: 003/124/SAP R11D 98-21
; CURRENT APPLICATION NUMBER: US/09/350,756
; EARLIER FILING DATE: 1999-07-09
; EARLIER FILING DATE: 1998-07-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 8
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Clostridium botulinum
; FEATURE:
US-09-350-756-8

Query Match
Best Local Similarity 1.6%; Score 51.6; DB 4; Length 1987;
Matches 151; Conservative 0; Mismatches 144; Indels 3; Gaps 1;

QY 2258 TTCTCAGTTCAATGAAGCAATATGAAGTATTTAACTCTTTCACCTACAGTCTTTCAGT 2317
DB 1011 TTCGAACTTAATGAGTCTCTATAAATAAAGCTATGATTAATAATAAATTTTGAATCA 1070
QY 2318 ATGCTATTTAAATTTACTTGGCCAGGTATAATTTGCCAGTCAGTCTCTTTATAGTGA 2377
DB 1071 ATGCTCTGTTTCATATTTTAATGAATCTATGATCCCTTATGGTGTAAACGGTGAAGA 1130
QY 2378 TTTATTTGGTTAGTATAATAAAT---ATTTTAACTAAATATAATAATCTATAATGTTAAA 2434
```

```
DB 1131 TTTTGTGCTAGTCTTAAAGATGCATTATTAAAGTATATATATATATATAGAGGAACTTT 1190
QY 2435 CATATGTTCAATAAAGCATAGCAGCTTTTGAATTAACATATATAAATAGCTCATATTTACA 2494
DB 1191 AATTTGTCAGTAGATAGATTAAAGATAAAGTTAATAATACACTTAGTACAGATATACC 1250
QY 2495 CTTACAGCTTTTTCATTTGATCAGTCTGAAATCTTTAGCACTTAAGGAAATGACTAT 2552
DB 1251 TTTTCAGCTTTTCCAATACGTAGATAAATCAAGATTATTATCTACATTTTACTGAATAT 1308

RESULT 13
US-09-255-829-7
; Sequence 7, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/782,893
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 2613 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2613
US-09-255-829-7

Query Match
Best Local Similarity 1.6%; Score 51.6; DB 3; Length 2613;
Matches 151; Conservative 0; Mismatches 144; Indels 3; Gaps 1;

QY 2258 TTCTCAGTTCAATGAAGCAATATGAAGTATTTAACTCTTTCACCTACAGTCTTTCAGT 2317
DB 2310 TTCGAACTTAATGAGTCTATAAATAAAGCTATGATTAATAATAAATTTTGAATCA 2369
QY 2318 ATGCTATTTAAATTTACTTGGCCAGGTATAATTTGCCAGTCAGTCTCTTTATAGTGA 2377
DB 2370 ATGCTCTGTTTTCATATTTTAATGAATTTCTATGATCCCTTATGGTGTAAACGGTGA 2429
```

Qy 2378 TTTATTGGTTAGTAATAATAAT---ATTTTAAACTAAATATAATAATCTATAATGTTAAA 2434
Db 2430 TTTTGTAGTCTTAAAGATGCAATTAATAAGTATATATATATAATAGAGAACTTT 2489
Qy 2435 CATATGTTCAATTAAGCATAGCACTTTGAAATTAACATATAATAAGTCTATTTTACA 2494
Db 2490 AATTGGTCAAGTAGATAGATTAAAGATAAAGTTAATAATACACTTAGTAGTACAGATATACC 2549
Qy 2495 CTTACAGCTTTTCAATTTGATCAGCTCTGAAATCTTTAGCACTTAAGGAAATGACTAT 2552
Db 2550 TTTTCAGCTTTTCCAAATACGTAGATAATCAAGAGATTATTCTACATTTACTGAATAT 2607

RESULT 14
US-09-255-829-1
; Sequence 1, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/782,893
; FILING DATE: 27-DEC-1996
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2616 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2616
; US-09-255-829-1

Query Match 1.6%; Score 51.6; DB 3; Length 2616;
Best Local Similarity 50.7%; Pred. No. 0.0051;
Matches 151; Conservative 0; Mismatches 144; Indels 3; Gaps 1;
Qy 2258 TTCTCAGTTCAATGAAGCAATGAAGTATTTAACCTTTTCACTACAGTTCTTGGCAAGT 2317
Db 2310 TTTCGAACACTTAAGTCTATATAAAGCTATGATTAAATAATAAATAAATTTTGAATCA 2369
Qy 2318 ATGCTATTTAAATTAATCTTGCCAGGTATTAATTCGCCAGTCAGTCTCTTTTATAGTGAGAAA 2377

Db 2370 ATGCTCTGTTTCATATTTAATGAATTCCTTATGATCCCTTATGTTAAACGGTTAGAAGA 2429
Qy 2378 TTTATTGGTTAGTAATAATAAT---ATTTTAAACTAAATATAATAATCTATAATGTTAAA 2434
Db 2430 TTTTGTAGTCTTAAAGATGCAATTAATAAGTATATATATATAATAGAGAACTTT 2489
Qy 2435 CATATGTTCAATTAAGCATAGCACTTTGAAATTAACATATAATAAGTCTATTTTACA 2494
Db 2490 AATTGGTCAAGTAGATAGATTAAAGATAAAGTTAATAATACACTTAGTAGTACAGATATACC 2549
Qy 2495 CTTACAGCTTTTCAATTTGATCAGCTCTGAAATCTTTAGCACTTAAGGAAATGACTAT 2552
Db 2550 TTTTCAGCTTTTCCAAATACGTAGATAATCAAGAGATTATTCTACATTTACTGAATAT 2607

RESULT 15
US-09-255-829-25
; Sequence 25, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/782,893
; FILING DATE: 27-DEC-1996
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2616 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2616
; US-09-255-829-25

Query Match 1.6%; Score 51.6; DB 3; Length 2616;
Best Local Similarity 50.7%; Pred. No. 0.0051;
Matches 151; Conservative 0; Mismatches 144; Indels 3; Gaps 1;
Qy 2258 TTCTCAGTTCAATGAAGCAATGAAGTATTTAACCTTTTCACTACAGTTCTTGGCAAGT 2317
Db 2310 TTTCGAACACTTAAGTCTATATAAAGCTATGATTAAATAATAAATAAATTTTGAATCA 2369

```
Qy 2318 ATGCTATTAAATTAAGGCGCAGGTATTAATGGCCAGTCAGTCTCTTTATAGTGAGAAA 2377
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 2370 ATGCTCTGTTTCATATTTAAATGAATTCATGATCCCTTATGGTGTAAACGGTTAGAAGA 2429
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 2378 TTTATTGGTTAGTAAATATAAT---ATTTTAACTAAATATATAATCTATATGTTTAAA 2434
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 2430 TTTTGATGCTAGTCTTAAAGATGCATTATTAAAGTATATATATGATAATAGAGGAACTTT 2489
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 2435 CATATGTTTCATTAAAGCATAGCAGCTTTGAAAATTAACTATATATAAATAGCTCATATTTACA 2494
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 2490 AATTGTCAGTAGATAGATTAAAGATAAAGTTAATAATACACTTAGTACAGATATACC 2549
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Qy 2495 CTTACAGCTTTTCATTGATCAGGTCGAAATCTTTAGCACTTAAGGAAAATGACTAT 2552
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 2550 TTTTCAGCTTTCCAAATACGTAGATAATCAAGATTATTATCTACATTTACTGAATAT 2607
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
```

Search completed: July 2, 2005, 19:50:47
Job time : 518 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 12:38:20 ; Search time 1635 seconds
(without alignments)
11607.762 Million cell updates/sec

Title: US-10-047-855-4
Perfect score: 3206
Sequence: 1 gtcgaccacgcgtccggc.....aaaaaaaaagggcgccgc 3206

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3206	100.0	3206	4 AAD06010	Aad06010 Human neu
2	3206	100.0	3206	8 ABS56723	AbS56723 Human NAR
3	3206	100.0	3206	12 ADJ57933	AdJ57933 Human NAR
4	3166	98.8	3499	5 ADL62788	AdL62788 Human ova
5	3166	98.8	5443	12 ADN05171	Adn05171 Antipsori
6	3166	98.8	5443	13 ADS74320	AdS74320 PRO83903
7	3164.4	98.7	3499	5 ABV29614	ABv29614 Human pro
8	3152.4	98.3	3498	5 ABV23741	ABv23741 Human pro
9	2481.8	77.4	2738	4 AAD06007	Aad06007 Human neu
10	2481.8	77.4	2738	12 ADJ57931	AdJ57931 Human NAR
11	2019	63.0	2019	13 ADR40144	ADR40144 Human NAR
12	1803	56.2	1803	4 AAH14905	AaH14905 Human CDN
13	1698.2	53.0	1929	4 AAH44173	AaH44173 Human gly
14	1471.2	45.9	3381	12 ADJ57938	AdJ57938 Rat NARC
15	1406	43.9	2477	5 ABA09667	ABa09667 Human bon
16	1376.4	42.9	1978	5 ABA09583	ABa09583 Human bon
17	991	30.9	2393	12 ADJ57947	AdJ57947 Rat NARC
18	633.4	19.8	768	4 AAH07671	AaH07671 Human CDN
19	486.4	15.2	546	4 AAS23890	AaS23890 Human ova
20	486.4	15.2	546	5 AAH82447	AaH82447 Human ova

c	21	469.4	14.6	471	6	ABV88538	ABv88538 Human col
	22	444.6	13.9	490	4	AAH12646	AaH12646 Human CDN
	23	427	13.3	427	9	ACH18980	Ach18980 Human adu
	24	414.8	12.9	520	5	ADL42779	AdL42779 Human ova
c	25	365.6	11.4	403	4	AAS24002	AaS24002 Human ova
	26	365.6	11.4	403	5	AAH82567	AaH82567 Human ova
c	27	363.4	11.3	396	5	ADL71235	AdL71235 Human ova
	28	363.4	11.3	396	5	ADL36395	AdL36395 Human ova
	29	326.6	10.2	425	5	ABV35865	ABv35865 Human pro
	30	326.6	10.2	425	5	ABV44927	ABv44927 Human pro
	31	326	10.2	327	5	ABV15063	ABv15063 Human pro
c	32	296.4	9.2	478	6	ABL81385	ABl81385 Human ova
	33	291	9.1	510	6	ABL79391	ABl79391 Human ova
c	34	285	8.9	298	2	AAT24062	AaT24062 Human gen
	35	271	8.5	359	5	ABV05894	ABv05894 Human pro
	36	265.2	8.3	474	4	AAK61234	AaK61234 Human imm
	37	236	7.4	14598	4	AAK80690	AaK80690 Human imm
	38	236	7.4	14598	4	AAK79627	AaK79627 Human imm
	39	208.4	6.5	463	4	AAI12461	AaI12461 Probe #23
	40	208.4	6.5	463	4	ABA54169	ABa54169 Human toe
	41	208.4	6.5	463	4	AAI33815	AaI33815 Probe #25
	42	208.4	6.5	463	4	ABA43713	ABa43713 Human bre
	43	208.4	6.5	463	4	ABA23914	ABa23914 Probe #23
	44	208.4	6.5	463	4	AAK27880	AaK27880 Human bon
	45	208.4	6.5	463	4	AAK02440	AaK02440 Human bra

ALIGNMENTS

RESULT 1
AAD06010
ID AAD06010 standard; DNA; 3206 BP.

AC AAD06010;

DT 31-JUL-2001 (first entry)

DE Human neuronal apoptosis regulated candidate (NARC) 16B DNA.

KW Human; neuronal apoptosis regulated candidate 16B; NARC 16B; cytostatic;
KW chromosome mapping; gene therapy; antisense therapy; lung disorder;
KW central nervous system disorder; apoptosis; spleen disorder; angina;
KW tuberculous; Goodpasture's syndrome; liver disorder; jaundice;
KW infectious disorder; brain disorder; cerebral oedema; gonorrhoea;
KW heart disorder; kidney disorder; glomerulonephritis; testes; virucide;
KW epididymis disorder; skeletal muscle disorder; pancreatic disorder;
KW diabetes; cytoprotectant; immunostimulant; tumour; antiinflammatory;
KW antimicrobial; neuroprotective; gynaecological; ds.

OS Homo sapiens.

FN WO200131007-A2.

XX PD 03-MAY-2001.

XX PF 20-OCT-2000; 2000WO-US029132.

XX PR 22-OCT-1999; 99US-0161188P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Chiang LW;

XX DR WPI; 2001-308641/32.

XX PT Rat brain polypeptides, nucleic acids and antibodies, useful for diagnosis and treatment of central nervous system disorders and disorders associated with aberrant apoptosis.

XX PS Claim 1; Page 148; 161pp; English.

XX CC The invention relates to human homologues of neuronal apoptosis regulated

CC candidate (NARC) nucleic acid molecules and proteins derived from rat
CC brain and programmed cell death libraries. The nucleic acids of the
CC invention are useful for assaying the presence of a nucleic acid molecule
CC and for chromosome mapping. They are also used in gene therapy and
CC antisense therapy. The NARC sequences are useful for treating central
CC nervous system disorders and disorders involving aberrant apoptosis, for
CC inducing an immune response and for isolating binding partners. Diseases
CC treated include spleen disorders (e.g. tuberculosis and congestive
CC splenomegaly), lung disorders (e.g. adult respiratory distress syndrome,
CC Goodpasture's syndrome and bronchial asthma), liver disorders (e.g.
CC jaundice and hepatic failure), infectious disorders (e.g. viral
CC hepatitis), brain disorders (e.g. cerebral oedema, hypertensive
CC encephalopathy and hydrocephalus), heart disorders (e.g. heart failure,
CC angina and myocardial infarction), kidney disorders (e.g. cysts and
CC glomerulonephritis), testes and epididymis disorders (e.g. gonorrhoea and
CC syphilis), skeletal muscle disorders (e.g. tumours) and pancreatic
CC disorders (e.g. pancreatitis and diabetes). The present sequence is human
CC neuronal apoptosis regulated candidate (NARC) 16B DNA
XX
SQ Sequence 3206 BP; 988 A; 580 C; 682 G; 956 T; 0 U; 0 Other;
Query Match 100.0%; Score 3206; DB 4; Length 3206;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCGACCCACGCGTCCGGCGAGGCA CGGACGGGGGGCGCCGGTACCTCTGCCCGCGGT 60
Db
QY 1 GTCGACCCACGCGTCCGGCGAGGCA CGGACGGGGGGCGCCGGTACCTCTGCCCGCGGT 60
Db
QY 61 CTTCTCCTCTCGGGCGGGGGCGGCGAGCGCGACTCGCGGACTAGCAACCCGGAGACG 120
Db
QY 61 CTTCTCCTCTCGGGCGGGGGCGGCGAGCGCGACTCGCGGACTAGCAACCCGGAGACG 120
Db
QY 121 ACATCATAAATAAATCCATCAGATGACACCTTCTCAGGTTGCTTTGAAATAAGAGGA 180
Db
QY 121 ACATCATAAATAAATCCATCAGATGACACCTTCTCAGGTTGCTTTGAAATAAGAGGA 180
Db
QY 181 ACTCTTTTACCAAGGAGAGTTTTCGGATATGTGGAAGCTGTGATGCTTTGGGAAACTGG 240
Db
QY 181 ACTCTTTTACCAAGGAGAGTTTTCGGATATGTGGAAGCTGTGATGCTTTGGGAAACTGG 240
Db
QY 241 AATCTCAAAATGCTGTGGCTCTTCTCCAGAGATGACAGGTTGAAGCATGCTATGG 300
Db
QY 241 AATCTCAAAATGCTGTGGCTCTTCTCCAGAGATGACAGGTTGAAGCATGCTATGG 300
Db
QY 301 AAAGCAACCATTTGACTCAGTAGAGAGTATCAGTTTCAGTATCGCTACTTCAAAGGGTAC 360
Db
QY 301 AAAGCAACCATTTGACTCAGTAGAGAGTATCAGTTTCAGTATCGCTACTTCAAAGGGTAC 360
Db
QY 361 TTTTGTAGAACCAAGACTATCGGTGCTCCATGTCAAGTGATAGTTTCAAGTGGGAGACT 420
Db
QY 361 TTTTGTAGAACCAAGACTATCGGTGCTCCATGTCAAGTGATAGTTTCAAGTGGGAGACT 420
Db
QY 421 CATCTCAACACGAGTCAATACCCCTTTAGAAAGGAAATTAATTATTGACGATGGACAA 480
Db
QY 421 CATCTCAACACGAGTCAATACCCCTTTAGAAAGGAAATTAATTATTGACGATGGACAA 480
Db
QY 481 TTTGGAAATCCAAATGCTGAACTCTGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCG 540
Db
QY 481 TTTGGAAATCCAAATGCTGAACTCTGGAATTCGGAATTCGGAATTCGGAATTCGGAATTCG 540
Db
QY 541 ATAAGATTACGTTTGCATTAATTCTGAAAAACCTCTCTGTGTCATTAACCAAGAAAAATTA 600
Db
QY 541 ATAAGATTACGTTTGCATTAATTCTGAAAAACCTCTCTGTGTCATTAACCAAGAAAAATTA 600
Db
QY 601 AAAAAATCTAGATTAGGTGAGTGCATGACATAGAAAGGCTCGAGGAGATGACCATGAT 660
Db
QY 601 AAAAAATCTAGATTAGGTGAGTGCATGACATAGAAAGGCTCGAGGAGATGACCATGAT 660
Db
QY 661 AGGGTATCTCCCACTGACTCCACAAATGCTCCAAATGCTGAGATATCTTAAATAGC 720
Db
QY 661 AGGGTATCTCCCACTGACTCCACAAATGCTCCAAATGCTGAGATATCTTAAATAGC 720
Db

QY 721 GACAAATGAGTTCAAGTGCAGGCAATTCACAGCCGGAGTGTTGTTATGGCTTCGACGCTGAT 780
Db
QY 721 GACAAATGAGTTCAAGTGCAGGCAATTCACAGCCGGAGTGTTGTTATGGCTTCGACGCTGAT 780
Db
QY 781 CTTTGGACAGAGTACAGCATACAGACGATGGAACAGATAACCTGGAATTAATCTTTGAT 840
Db
QY 781 CTTTGGACAGAGTACAGCATACAGACGATGGAACAGATAACCTGGAATTAATCTTTGAT 840
Db
QY 841 TTTTTCGAAAGAGATCTCAGTGAGCAGCATGTAGTTTCAGGGTGATGCCCTTCTCTGGACATGTG 900
Db
QY 841 TTTTTCGAAAGAGATCTCAGTGAGCAGCATGTAGTTTCAGGGTGATGCCCTTCTCTGGACATGTG 900
Db
QY 901 GGTACAGCTTGTCTCTTATATCCCAATTCCTGAGTGGGAAAGAGTCTGGAATTTCTT 960
Db
QY 901 GGTACAGCTTGTCTCTTATATCCCAATTCCTGAGTGGGAAAGAGTCTGGAATTTCTT 960
Db
QY 961 ACTTTTCCCATCATGAGCAGAAATTTCCCGGAAAAACAATAGGCAAAAGTGGAGCTTGCATAT 1020
Db
QY 961 ACTTTTCCCATCATGAGCAGAAATTTCCCGGAAAAACAATAGGCAAAAGTGGAGCTTGCATAT 1020
Db
QY 1021 ATAAATTTAAGCCATTACCAGGATACAGTTTGTGACATGAAATCTTCAATTTTCCAAGTAT 1080
Db
QY 1021 ATAAATTTAAGCCATTACCAGGATACAGTTTGTGACATGAAATCTTCAATTTTCCAAGTAT 1080
Db
QY 1081 TGGAAAGCCAAAGAAATACCATTTGGATGTTGCCCATCGAGGTGCAGGAAACTCTCAACAAT 1140
Db
QY 1081 TGGAAAGCCAAAGAAATACCATTTGGATGTTGCCCATCGAGGTGCAGGAAACTCTCAACAAT 1140
Db
QY 1141 GCCACGCTGGCTAAAGTTCAAGAAATACTATTGCTCTTTTAAAGAAATGCTGCTAGTCAAT 1200
Db
QY 1141 GCCACGCTGGCTAAAGTTCAAGAAATACTATTGCTCTTTTAAAGAAATGCTGCTAGTCAAT 1200
Db
QY 1201 GGTGCGACCTTTGTAGAAATTTGACGTTACACCTTTTCAAGGACTTTTGTGCGCGGTGATAT 1260
Db
QY 1201 GGTGCGACCTTTGTAGAAATTTGACGTTACACCTTTTCAAGGACTTTTGTGCGCGGTGATAT 1260
Db
QY 1261 CATGATCTTACCTGTTGTTTGTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTA 1320
Db
QY 1261 CATGATCTTACCTGTTGTTTGTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTA 1320
Db
QY 1321 TTTGAAATTCAGTAAAAAGAAATTAACATTTGACAACTCCAGTTGTTAAAGCTCACTCAT 1380
Db
QY 1321 TTTGAAATTCAGTAAAAAGAAATTAACATTTGACAACTCCAGTTGTTAAAGCTCACTCAT 1380
Db
QY 1381 GTGACTGCACTGAAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTT 1440
Db
QY 1381 GTGACTGCACTGAAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTT 1440
Db
QY 1441 TCAGAAATCAGCCATTTCTTCTTAAAGATGGTTTTAGAGTCTTTTGCAGAAAGATGA 1500
Db
QY 1441 TCAGAAATCAGCCATTTCTTCTTAAAGATGGTTTTAGAGTCTTTTGCAGAAAGATGA 1500
Db
QY 1501 GGGTTTAACTGAAATTAAGATGATCTGCCAGAAAGGATGGAATGTTGGGATGGTAAC 1560
Db
QY 1501 GGGTTTAACTGAAATTAAGATGATCTGCCAGAAAGGATGGAATGTTGGGATGGTAAC 1560
Db
QY 1561 TTATCAACATATTTTGACATGAATCTGTTTTTGGATATAATTTTAAACCTGTTTGTAGAA 1620
Db
QY 1561 TTATCAACATATTTTGACATGAATCTGTTTTTGGATATAATTTTAAACCTGTTTGTAGAA 1620
Db
QY 1621 AATTTCTGGGAGGAGGAAATAGTGTCTTCTTCAATTTGATGAGATATTTGCACAAATGGTT 1680
Db
QY 1621 AATTTCTGGGAGGAGGAAATAGTGTCTTCTTCAATTTGATGAGATATTTGCACAAATGGTT 1680
Db
QY 1681 CGGCAAAAGCAGAAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTAT 1740
Db
QY 1681 CGGCAAAAGCAGAAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTAT 1740
Db
QY 1741 CCTGAACTCATGGAACCTCAGATCTCGGACAAACCCCAATTCGAATGAGCTTTGACAGCTTT 1800
Db
QY 1741 CCTGAACTCATGGAACCTCAGATCTCGGACAAACCCCAATTCGAATGAGCTTTGACAGCTTT 1800
Db
QY 1801 GAAATCTACTGGGGATAAATGTATACATCTGAAGACTTGTCTCAGAAAAACCCATCTATATT 1860

1801 GAAATCTACTGGGTAATGATACATCTGAGACTTCTCAGAAACCCATCTTATATT 1860
1861 CAAGAGCAAAAGCTAAGGAGCTAGTCATATCTCTGCTGGGGTGATGATACCAATGATCT 1920
1861 CAAGAGCAAAAGCTAAGGAGCTAGTCATATCTCTGCTGGGGTGATGATACCAATGATCT 1920
1921 GAAAACAGAGGAAATTTGAAGGAACTTGGAGTTAATGGTCTAATTTATGATAGGATATAT 1980
1921 GAAAACAGAGGAAATTTGAAGGAACTTGGAGTTAATGGTCTAATTTATGATAGGATATAT 1980
1981 GATTGATGCTGCAACCAACCAATATATTCAGTGGAGCAATTCGAAAGCCCTGAGCAG 2040
1981 GATTGATGCTGCAACCAACCAATATATTCAGTGGAGCAATTCGAAAGCCCTGAGCAG 2040
2041 GAATTCGCAGAGCTTAAAGAGCTTGTGTGCTCCACCTGTTAGCGCTTGTCCCTCATCT 2100
2041 GAATTCGCAGAGCTTAAAGAGCTTGTGTGCTCCACCTGTTAGCGCTTGTCCCTCATCT 2100
2101 TTGTGTGGGGAGTCTGATATCCATGTTGATGCCAGCAATGATTAACGTGAGAAATGCT 2160
2101 TTGTGTGGGGAGTCTGATATCCATGTTGATGCCAGCAATGATTAACGTGAGAAATGCT 2160
2161 TAGTTTTTATTCACACAGAGCTCAATTTGGGGGGTGCACCGCTGTTCTGGGTATTCATTT 2220
2161 TAGTTTTTATTCACACAGAGCTCAATTTGGGGGGTGCACCGCTGTTCTGGGTATTCATTT 2220
2221 TTTCATCAGTGAAGCTTGTGATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAATAA 2280
2221 TTTCATCAGTGAAGCTTGTGATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAATAA 2280
2281 TGAAGTATTTAACTCTTTCACTACAGTCTTGCAGATGATGCTATTTAAATTTACTTGGCCA 2340
2281 TGAAGTATTTAACTCTTTCACTACAGTCTTGCAGATGATGCTATTTAAATTTACTTGGCCA 2340
2341 GGTATAATTTGCAGTCAGTCTCTTTATAGTGAGAAAAATTTATTTGGTGTAGTATATAAATA 2400
2341 GGTATAATTTGCAGTCAGTCTCTTTATAGTGAGAAAAATTTATTTGGTGTAGTATATAAATA 2400
2401 TTTTAACTTAAATATATAATCTATTAATCTTAAACATATGTTTCATTAAGCAATAGCACT 2460
2401 TTTTAACTTAAATATATAATCTATTAATCTTAAACATATGTTTCATTAAGCAATAGCACT 2460
2461 TTGAAATTAACATATATAATAGTCTCATATTTACACTACAGCTTTTCAITTTGATCAGGTC 2520
2461 TTGAAATTAACATATATAATAGTCTCATATTTACACTACAGCTTTTCAITTTGATCAGGTC 2520
2521 TGAATCTTTAGCACTTAAGGAAATGACTATGCTATGCTAATTTACCTGACCATGAAAAAA 2580
2521 TGAATCTTTAGCACTTAAGGAAATGACTATGCTAATTTACCTGACCATGAAAAAA 2580
2581 TAAGTACCTCAATGATCATCTTTGACCTGGTGTATTCACCACTGCAAAATCTTTTGCCCA 2640
2581 TAAGTACCTCAATGATCATCTTTGACCTGGTGTATTCACCACTGCAAAATCTTTTGCCCA 2640
2641 TCTTGATATAGGTATTTTTTACATGGGTTGACATGCAACCAACCACTTTTCAITTCAGT 2700
2641 TCTTGATATAGGTATTTTTTACATGGGTTGACATGCAACCAACCACTTTTCAITTCAGT 2700
2701 ATGAACCTTGAGGCTGCTGCCATTTTTCATTAACCAACCAAGCTGAGGTGAACCTTC 2760
2701 ATGAACCTTGAGGCTGCTGCCATTTTTCATTAACCAACCAAGCTGAGGTGAACCTTC 2760
2761 GAACTTGTTCATTAATCTTTTCAAAAGTGTGTTTACATCAATGTTAAATTTCAAAATG 2820
2761 GAACTTGTTCATTAATCTTTTCAAAAGTGTGTTTACATCAATGTTAAATTTCAAAATG 2820
2821 CTGCAGGGTAAATTTATGATATAAATATTTAGTAAGAAAAAGTATGTTATGTCATCTAGT 2880
2821 CTGCAGGGTAAATTTATGATATAAATATTTAGTAAGAAAAAGTATGTTATGTCATCTAGT 2880
2881 AGAATAGATCAACAATACAAATTCATTTCAAGTCAATGCTTTAGGTGTTAAGCATGAGAT 2940

2881 AGAATAGATCAACAATACAAATTCATTTCAAGTCAATGCTTTAGGTGTTAAGCATGAGAT 2940
2941 TGTACATGTTTACTGTTAGGTCCTTGCATCTGTGCTGCTAGGTGATGAGAAATGTC 3000
2941 TGTACATGTTTACTGTTAGGTCCTTGCATCTGTGCTGCTAGGTGATGAGAAATGTC 3000
3001 AAGGACTGGAGTATTTGTTGCTTAAATAAAAAAGGCTGTTTGTAGGCGTTTAAATAT 3060
3001 AAGGACTGGAGTATTTGTTGCTTAAATAAAAAAGGCTGTTTGTAGGCGTTTAAATAT 3060
3061 GCTTATTTTGTGTCCTCTCACTACTATTACACTGTTGCTTTGCTGGTGTGTTTGT 3120
3061 GCTTATTTTGTGTCCTCTCACTACTATTACACTGTTGCTTTGCTGGTGTGTTTGT 3120
3121 ATGTCGCTGTGTTATACACTAGTATTAATTTCCATGCAGAAAAATAAATGTCCTGAATCT 3180
3121 ATGTCGCTGTGTTATACACTAGTATTAATTTCCATGCAGAAAAATAAATGTCCTGAATCT 3180
3181 CAAAAAATAAAAAAAGGCGCGCGC 3206
3181 CAAAAAATAAAAAAAGGCGCGCGC 3206

RESULT 2
ABS56723
ID ABS56723 standard; cDNA; 3206 BP.
XX
AC ABS56723;
XX
DT 24-FEB-2003 (first entry)
DE Human NARC16 cDNA.
XX
KW NARC16; cytostatic; immunosuppressive; dermatological; cardiac;
KW antiinflammatory; nephrotropic; anti-HIV; nontropic; neuroprotective;
KW antianemic; cerebroprotective; vasotropic; antidiabetic; anticonvulsant;
KW immunosuppressive; thymimetic; immunostimulant; antianemic; hypotensive;
KW tranquiliser; neuroleptic; gene therapy; gene mapping; apoptosis;
KW AIDS; cell cycle disruption; programmed cell death regulation;
KW viral infection; nucleosome assembly; phosphate homeostasis;
KW cell cycle regulation; cancer; follicular lymphoma; carcinoma;
KW p53 mutation; graft rejection; hormone-dependent tumour;
KW autoimmune disorder; valvular heart disease;
KW systemic lupus erythematosus; diabetes; Hashimoto's thyroiditis;
KW immune-mediated glomerulonephritis; virus-induced lymphocyte depletion;
KW acquired immunodeficiency syndrome; neurodegenerative disease; stroke;
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW spinal muscular atrophy; retinitis pigmentosa; cerebellar degeneration;
KW myelodysplastic syndrome; ischaemic injury; myocardial infarction;
KW reperfusion injury; liver disease; idiopathic dilated cardiomyopathy;
KW ischaemic cardiomyopathy; aplastic anaemia; chronic neutropenia; mania;
KW myelodysplastic syndrome; central nervous system disorder; anxiety;
KW senile dementia; Huntington's disease; hypertension; schizophrenia;
KW severe bipolar affective disorder; human; chromosome 4q11-4q21; ss.

XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 145..2163
XX FT /*tag= a
XX FT /product= "NARC16"
XX
XX PN WO200281516-A2.
XX
XX PD 17-OCT-2002.
XX
XX PF 16-JAN-2002; 2002WO-US001098.
XX
XX PR 16-JAN-2001; 2001US-0262306P.
XX PR 15-JAN-2002; 2002US-00047855.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX

PI	Chiang LW;
DR	WPI; 2003-058503/05.
XX	P-PSDB; ABB84606.
DR	
XX	Novel isolated programmed cell death-related polypeptide, NARC10 and
PT	NARC16, useful for treating disorders associated with abnormal apoptotic
PT	process e.g. Alzheimer's disease, cancer, myocardial infarction, stroke.
XX	
PS	Claim 1; Fig 4A-C; 123pp; English.
XX	
CC	This invention describes novel cell death-related polypeptides NARC10 and
CC	NARC16, located on chromosome 4q11-q21 and which have cardiant, antiHIV,
CC	immunosuppressive, dermatological, antiinflammatory, cerebroprotective,
CC	nootropic, neuroprotective, antianemic, cardiant, vasotropic, antianemic,
CC	antidiabetic, immunosuppressive, cytoskeletal; thyromimetic, nephrotropic,
CC	immunostimulant, anticonvulsant, tranquiliser, hypotensive and
CC	neuroleptic activity and can be used in gene therapy. The products of the
CC	invention can be used to modulate NARC10 or NARC16 polypeptides or
CC	polynucleotides, to map NARC genes on a chromosome, e.g. to locate gene
CC	regions associated with genetic disease or to associate NARC10 or NARC16
CC	with a disease. The polypeptides are also useful for modulating the
CC	apoptotic process, and are therefore useful for modulating, and treating
CC	disorders associated with increased apoptosis, inhibition of apoptosis or
CC	disruptions in cell cycle, for regulating cellular functions including
CC	programmed cell death, nucleosome assembly, phosphate homeostasis and the
CC	cell cycle. Preferably, the products of the invention are useful for
CC	treating disorders associated with abnormally low rate or abnormally high
CC	rate of apoptosis e.g. cancers including follicular lymphomas, carcinomas
CC	with p53 mutations, or hormone-dependent tumours, autoimmune disorders
CC	including systemic lupus erythematosus, diabetes, graft rejection,
CC	Hashimoto's thyroiditis and immune-mediated glomerulonephritis and viral
CC	infections e.g. infections caused by herpes viruses, virus-induced
CC	lymphocyte depletion (including acquired immunodeficiency syndrome
CC	(AIDS)), neurodegenerative diseases manifested by loss of specific sets
CC	of neurons (including Alzheimer's disease, Parkinson's disease,
CC	amyotrophic lateral sclerosis, spinal muscular atrophy, retinitis
CC	pigmentosa, and cerebellar degeneration), myelodysplastic syndromes
CC	(including aplastic anaemia), ischaemic injuries (including myocardial
CC	infarction, stroke and reperfusion injury) and toxin (e.g. alcohol)
CC	induced liver disease, idiopathic dilated cardiomyopathy, ischaemic
CC	cardiomyopathy and valvular heart disease, aplastic anaemia, chronic
CC	neutropenia, and myelodysplastic syndromes, central nervous system
CC	disorders, senile dementia, Huntington's disease, hypertension,
CC	schizophrenia, attention deficit disorder, mania, anxiety, severe bipolar
CC	affective disorder (BP-I). This sequence encodes the human NARC16 protein
CC	described in the method of the invention
XX	
SQ	Sequence 3206 BP; 988 A; 580 C; 682 G; 956 T; 0 U; 0 Other;
	Query Match 100.0%; Score 3206; DB 8; Length 3206;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 3206; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 GTGCACCCAGCGCTCGGGCGGAGGCA CGGACGGCGGGGCCCGGTACTCTTCGCCGCGGT 60
DB	1 GTGCACCCAGCGCTTCGGGCGAGGACG GACGGCGGGGCCCGGTACTCTTCGCCGCGGT 60
QY	61 CCTCGCTCTCGGGCGGGCGGGCGGACG CGGACTGCGGA CTAGCGAACCCGAGCACG 120
DB	61 CCTCGCTCTCGGGCGGGCGGGCGGACG CGGACCTGCGGACTAGCGAA CCCGAGCACG 120
QY	121 ACATCATATAAATTAATTCATCAGAATGACACC TTCTCAGGTGCTTTTGAAAATAAGAGGA 180
DB	121 ACATCATATAAATTAATTCATCAGAATGAC ACCCTTCTCAGGTGCTTTTGAAAATAAGAGGA 180
QY	181 ACTCTTTTACCAGSAGAAGTTTTTGGGATATGTGGAAGCTGTGATGCTTTTGGGAAA CTCG 240
DB	181 ACTCTTTTACAGSAGAAGTTTTTGGGATATGTGGAAGCTGTGATGCTTTTGGGAAA CTCG 240
QY	241 AANTCCTCAAATTCGTGTGGCTCTTTTCCAGAGATGACACAGGTGGAAGCATGCTATGG 300
DB	241 AATCTCTTAATATCTGTGGTCTTTCTTCCAGAGATGACACAGGTGGAAGCATGCTATGG 300

PN XX US2004009553-A1.
PD XX 15-JAN-2004.
XX XX 30-APR-2003; 2003US-00426776.
XX XX 27-SEP-1999; 99US-00406045.
PR 22-OCT-1999; 99US-0161188P.
XX 31-JAN-2000; 2000US-00495823.
PR 28-FEB-2000; 2000US-0185517P.
XX 20-OCT-2000; 2000US-00692785.
PR 31-JAN-2001; 2001US-00773426.
XX 28-FEB-2001; 2001US-00795691.
PR 31-OCT-2001; 2001US-0335003P.
XX 31-OCT-2001; 2001US-0335037P.
PR 25-MAR-2002; 2002US-00105992.
XX 28-AUG-2002; 2002US-00229662.
PR 30-OCT-2002; 2002US-00284014.
XX 30-OCT-2002; 2002US-00284059.
PR 09-DEC-2002; 2002US-00314881.
XX (MILL-) MILLENNIUM PHARM INC.
XX Glucksmann MA, Williamson MJ, Tsai F, Rudolph-Owen LA;
PI Kapeller-Libermann R, Meyers RE, Chiang LW, Hunter JJ;
XX WPI; 2004-090469/09.
XX New isolated nucleic acid molecules and polypeptides (e.g. 27411, 23413,
PT 22438, 23553, NARC SCL or NARC 1) useful for diagnosing, preventing or
PT treating disorders associated with the protein, e.g. cancer,
PT atherosclerosis or AIDS.
XX Claim 1; SEQ ID NO 34; 260pp; English.
XX The present invention provides isolated nucleic acid molecules and
CC proteins designated 27411, 23413, 22438, 23553, 25278, 26212, NARC SCL,
CC NARC 10A, NARC 1, NARC 12, NARC 13, NARC 17, NARC 25, NARC 3, NARC 4,
CC NARC 7, NARC 8, NARC 11, NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20,
CC NARC 26, NARC 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C,
CC NARC 8B, NARC 9, NARC 2A, NARC 16B, NARC 1C, NARC 1A, NARC 25, 86604 or
CC 32222. The invention is useful in diagnosing, preventing or treating
CC disorders such as cancer, lung diseases, cirrhosis, hepatitis,
CC atherosclerosis, myocardial infarction, inflammation, anaemia,
CC glomerulonephritis, osteoporosis, AIDS (acquired immunodeficiency
CC syndrome), Parkinson's disease, Alzheimer's disease, stroke and
CC dermatitis. These may also be used in drug screening. The invention is
CC also useful in gene therapy. The present sequence is human neuronal
CC apoptosis regulated candidate (NARC) cDNA.
XX
SQ Sequence 3206 BP; 988 A; 580 C; 682 G; 956 T; 0 U; 0 Other;
Query Match 100.0%; Score 3206; DB 12; Length 3206;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCGACCCACGGCTCGCGCGAGGACGACGCGCGCGCGCGCGCTCTGCGCGCGGT 60
DB 1 GTCGACCCACGGCTCGCGCGAGGACGACGCGCGCGCGCGCGCTCTGCGCGCGGT 60
QY 61 CCTCGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTGCGGACCTAGCGAACCCGAGACG 120
DB 61 CCTCGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTGCGGACCTAGCGAACCCGAGACG 120
QY 121 ACATCATATAAATAATCCATCAGATGACACCTTCTCAGGTTGCGCTTTGAAATAGAGGA 180
DB 121 ACATCATATAAATAATCCATCAGATGACACCTTCTCAGGTTGCGCTTTGAAATAGAGGA 180
QY 181 ACTCTTTTACAGGAGAGGTTTTTGGATATGTTGGAAGCTGTGATGCTTTGGGAACTGG 240
DB 181 ACTCTTTTACAGGAGAGGTTTTTGGATATGTTGGAAGCTGTGATGCTTTGGGAACTGG 240
QY 241 AATCCTCAAAATGCTGTGGCTCTTCTTCCAGAGAAATGACACAGGTGAAAGCATGCTATGG 300

DB 241 AATCCTCAAAATGCTGTGGCTCTTCTTCCAGAGAAATGACACAGGTGAAAGCATGCTATGG 300
QY 301 AAAGCAACCATTTGTAAGTATGAGGAGTATCAGTTTCAAGTATGCTTCAAGAGGTAC 360
DB 301 AAAGCAACCATTTGTAAGTATGAGGAGTATCAGTTTCAAGTATGCTTCAAGAGGTAC 360
QY 361 TTTTGTAGAACCAAGACTATCGGTGCTTCAAGTATGCTTCAAGTATGCTTCAAGTATGCT 420
DB 361 TTTTGTAGAACCAAGACTATCGGTGCTTCAAGTATGCTTCAAGTATGCTTCAAGTATGCT 420
QY 421 CATCTACAAACCAAGTATCAATTAACCCCTTTAGAAAGCAAAATTTATTTAGCAGTACAAA 480
DB 421 CATCTACAAACCAAGTATCAATTAACCCCTTTAGAAAGCAAAATTTATTTAGCAGTACAAA 480
QY 481 TTTTGGAAATCCAAATGCTTGAACCTCTGGAATCTGGAATGCTGATGCTGATGCTGATG 540
DB 481 TTTTGGAAATCCAAATGCTTGAACCTCTGGAATCTGGAATGCTGATGCTGATGCTGATG 540
QY 541 ATAAGATTAGCTTTGCAATTTCTGAAAAACCTCTGTTGTTCAATAACCAAGAAAAATTA 600
DB 541 ATAAGATTAGCTTTGCAATTTCTGAAAAACCTCTGTTGTTCAATAACCAAGAAAAATTA 600
QY 601 AAAAAATCTAGATTTAGGGTGAAGCTGACACTAGAAAGGCTGGAGGAAGATGACGATGAT 660
DB 601 AAAAAATCTAGATTTAGGGTGAAGCTGACACTAGAAAGGCTGGAGGAAGATGACGATGAT 660
QY 661 AGGGTATCTCCACTGTAATCTGCAAAAATGTTCAATAGCTTGGAGATATCCTTTAAATG 720
DB 661 AGGGTATCTCCACTGTAATCTGCAAAAATGTTCAATAGCTTGGAGATATCCTTTAAATG 720
QY 721 GACAAATGAGTTCAAGTGCAGGCAATTCACAGCCGGAGTGTGTTATGGCTTGCAGCTGAT 780
DB 721 GACAAATGAGTTCAAGTGCAGGCAATTCACAGCCGGAGTGTGTTATGGCTTGCAGCTGAT 780
QY 781 CGTTGGACAGATACAGCATACAGAGTGAACACAGATAACCTGGAACCTTAATCTTTGAT 840
DB 781 CGTTGGACAGATACAGCATACAGAGTGAACACAGATAACCTGGAACCTTAATCTTTGAT 840
QY 841 TTTTTCGAAAGAAATCTCAGTGCAGCAGTATGTTTCAAGGAGTATGCTTCTGACATGTTG 900
DB 841 TTTTTCGAAAGAAATCTCAGTGCAGCAGTATGTTTCAAGGAGTATGCTTCTGACATGTTG 900
QY 901 GGTACAGCTTGTCTCTTATCATCCACATGCTGAGAGTGAAGAGTGTGGAATCTTT 960
DB 901 GGTACAGCTTGTCTCTTATCATCCACATGCTGAGAGTGAAGAGTGTGGAATCTTT 960
QY 961 ACTCTTCCCATCATGAGCAGAAATTCGCGGAAACCAATAGGCAAAAGTGAAGTGAATAT 1020
DB 961 ACTCTTCCCATCATGAGCAGAAATTCGCGGAAACCAATAGGCAAAAGTGAAGTGAATAT 1020
QY 1021 ATAATTATTAAGCCATTAACAGGATACAGTTGTGACATGAATAATCTTCAATTTCAAGTAT 1080
DB 1021 ATAATTATTAAGCCATTAACAGGATACAGTTGTGACATGAATAATCTTCAATTTCAAGTAT 1080
QY 1081 TGGAGCCAAAGAAATACCATTTGGGATTTGGCCATCGAGGTGAGGAAACTCTACAAACAT 1140
DB 1081 TGGAGCCAAAGAAATACCATTTGGGATTTGGCCATCGAGGTGAGGAAACTCTACAAACAT 1140
QY 1141 GCCCAGCTGGCTAAAGTTCAAGAAATATCTATTTCTTTTAAAGAAATGCTGTAGTATCAT 1200
DB 1141 GCCCAGCTGGCTAAAGTTCAAGAAATATCTATTTCTTTTAAAGAAATGCTGTAGTATCAT 1200
QY 1201 GGTGACGCTTTGTGAGAAATTTGACGTACACTTTCAAAGGACTTTGTGCGCGGTATAT 1260
DB 1201 GGTGACGCTTTGTGAGAAATTTGACGTACACTTTCAAAGGACTTTGTGCGCGGTATAT 1260
QY 1261 CATGATCTTACCTGTTGTTGATCATCAAAAAGAAATTTGATGCTGATGCTGATGCTGAT 1320
DB 1261 CATGATCTTACCTGTTGTTGATCATCAAAAAGAAATTTGATGCTGATGCTGATGCTGAT 1320
QY 1321 TTTTCAAAATCCAGTAAAGAAATTAACATTTGACCACTCCAGTTGTTTAAAGCTCACTCAT 1380
DB 1321 TTTTCAAAATCCAGTAAAGAAATTAACATTTGACCACTCCAGTTGTTTAAAGCTCACTCAT 1380

1321 TTGAAATTCAGTAAAGAAATTAACATTTGACCAACTCAGTTGTTAAAGCTCACTCAT 1380
1381 GTGACTGCATGAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTT 1440
1381 GTGACTGCATGAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTT 1440
1441 TCAGAAATCAGCCATTCCTCTCTTTAAGATGGTTTGAAGTCTTTGCGAGAAGATGTA 1500
1441 TCAGAAATCAGCCATTCCTCTCTTTAAGATGGTTTGAAGTCTTTGCGAGAAGATGTA 1500
1501 GGGTTTAACTTGAATTAATAATGATCTGCAGCAAGGATGGAATGCGGATGTTAAC 1560
1501 GGGTTTAACTTGAATTAATAATGATCTGCAGCAAGGATGGAATGCGGATGTTAAC 1560
1561 TTATCAACATATTTTGAATGAATCTGTTTTCGATATATTTTAAAGAACTGTTTAGAA 1620
1561 TTATCAACATATTTTGAATGAATCTGTTTTCGATATATTTTAAAGAACTGTTTAGAA 1620
1621 AATTCTGGGAGGAGGAAATAGTGTCTTCTTCAATTTGATGCGATATTTGCAATGGTT 1680
1621 AATTCTGGGAGGAGGAAATAGTGTCTTCTTCAATTTGATGCGATATTTGCAATGGTT 1680
1681 CGCAAAAGCAGAACTATCCGATCTATTTTAACTCAAGGAAATCTGAGATTTAT 1740
1681 CGCAAAAGCAGAACTATCCGATCTATTTTAACTCAAGGAAATCTGAGATTTAT 1740
1741 CCTGAACTCATGGAACCTCAGATCTCGGACAAACCCCAATTCGATGAGCTTTGACAGTTT 1800
1741 CCTGAACTCATGGAACCTCAGATCTCGGACAAACCCCAATTCGATGAGCTTTGACAGTTT 1800
1801 GAAATCTACTGGGGATTAATGATACATACGAAAGACTTTGCTCAGAAACCCCACTATAT 1860
1801 GAAATCTACTGGGGATTAATGATACATACGAAAGACTTTGCTCAGAAACCCCACTATAT 1860
1861 CAAGAGCAAAAGCTAAGGAGCTAGTCAATATCTGCTGGGGATGATACCAATGATCT 1920
1861 CAAGAGCAAAAGCTAAGGAGCTAGTCAATATCTGCTGGGGATGATACCAATGATCT 1920
1921 GAAACAGAGGAAATTTGAAGGAACTTGGAGTTAATGGTCTAATTTATGATAGGATATAT 1980
1921 GAAACAGAGGAAATTTGAAGGAACTTGGAGTTAATGGTCTAATTTATGATAGGATATAT 1980
1981 GATTGATGCTGAAACCAATATATTTCCAGAGTGAGCAATTTGAAACGCTGAGGAG 2040
1981 GATTGATGCTGAAACCAATATATTTCCAGAGTGAGCAATTTGAAACGCTGAGGAG 2040
2041 GAATTCAGAGCTTAAAGAGCTGTTTGTGTCACAGTGTAGCGCTTTGCTCCCTCATCT 2100
2041 GAATTCAGAGCTTAAAGAGCTGTTTGTGTCACAGTGTAGCGCTTTGCTCCCTCATCT 2100
2101 TTGTGGGGAGTCTGATATCCATGTTGGATGCCAAGCAATTCATTAAGGATGCT 2160
2101 TTGTGGGGAGTCTGATATCCATGTTGGATGCCAAGCAATTCATTAAGGATGCT 2160
2161 TAGTTTTTATGACAGAGTCAATTTTGGGGGCTGCAACGCTGTTCTGGGTATTCATTT 2220
2161 TAGTTTTTATGACAGAGTCAATTTTGGGGGCTGCAACGCTGTTCTGGGTATTCATTT 2220
2221 TTCATCAGTACGATGTTGATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAATTA 2280
2221 TTCATCAGTACGATGTTGATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAATTA 2280
2281 TGAAGTATTTAACTCTTTTCACTACAGTCTTTCGAGTATGCTATTTAAATTAAGTGGCA 2340
2281 TGAAGTATTTAACTCTTTTCACTACAGTCTTTCGAGTATGCTATTTAAATTAAGTGGCA 2340
2341 GGTATTAATTTCCAGTCACTCTTTTATAGTGAGAAATTTATTTGGTTAGTAAATAAATA 2400
2341 GGTATTAATTTCCAGTCACTCTTTTATAGTGAGAAATTTATTTGGTTAGTAAATAAATA 2400
2401 TTTTAACTTAAATATATATCTATTAATGTTTAAACATATGTTTCAATTAAGCATAGCACT 2460
2401 TTTTAACTTAAATATATATCTATTAATGTTTAAACATATGTTTCAATTAAGCATAGCACT 2460

QY 2461 TTGAAATTAACATATATAAATAGCTCATATTTACACTTACAGCTTTTCACTTATGATCAGGTC 2520
Db 2461 TTGAAATTAACATATATAAATAGCTCATATTTACACTTACAGCTTTTCACTTATGATCAGGTC 2520
QY 2521 TGAATCTTTTAGCCTTAAGGAAATGATATGCAATAATATATCTGACCATGAAATGAAAA 2580
Db 2521 TGAATCTTTTAGCCTTAAGGAAATGATATGCAATAATATATCTGACCATGAAATGAAAA 2580
QY 2581 TAACTACCTCAATGATGCTTTCACCTGGTATTTCCAACTGCAACAAATCTTCTGCGCA 2640
Db 2581 TAACTACCTCAATGATGCTTTCACCTGGTATTTCCAACTGCAACAAATCTTCTGCGCA 2640
QY 2641 TCTTGTATATAGTATTTTACATGGTGTGACATGCGACACAAACCACTTTTCACTTCACT 2700
Db 2641 TCTTGTATATAGTATTTTACATGGTGTGACATGCGACACAAACCACTTTTCACTTCACT 2700
QY 2701 ATGAACCTTGAAGCTGCTGCCAATTTTCCACTTAAACCAACCCAGCTGAAAGTGAACCTC 2760
Db 2701 ATGAACCTTGAAGCTGCTGCCAATTTTCCACTTAAACCAACCCAGCTGAAAGTGAACCTC 2760
QY 2761 GAACTCTGTTTCATAAATCTTTTCAAAAGTGTGTTTACATCAATGTTAAATTTCAAAATG 2820
Db 2761 GAACTCTGTTTCATAAATCTTTTCAAAAGTGTGTTTACATCAATGTTAAATTTCAAAATG 2820
QY 2821 CTGACGGGTAAATTTAATGATATAAATATTTAGTAAAGAAAGTATGTTTGCATACCTTACT 2880
Db 2821 CTGACGGGTAAATTTAATGATATAAATATTTAGTAAAGAAAGTATGTTTGCATACCTTACT 2880
QY 2881 AGAATAGATCAACATACAAATCAATTCAGTGCATGCTTTAGGTGTTAAGCATGAGAT 2940
Db 2881 AGAATAGATCAACATACAAATCAATTCAGTGCATGCTTTAGGTGTTAAGCATGAGAT 2940
QY 2941 TGTACATGTTTACTGTTAGTCTTGCATCTGCTGCTAGTGCATGATGAGAGATGTC 3000
Db 2941 TGTACATGTTTACTGTTAGTCTTGCATCTGCTGCTAGTGCATGATGAGAGATGTC 3000
QY 3001 AAGGACTGAGCTATTTTGTGCTTAAAGGCTGTTTGTAGGCGTTTAAATAT 3060
Db 3001 AAGGACTGAGCTATTTTGTGCTTAAAGGCTGTTTGTAGGCGTTTAAATAT 3060
QY 3061 GCTTATTTTGTGCTCTCTACCTACCTATTAACACATGCTGCTTGTGGGTTTGTGTTTGT 3120
Db 3061 GCTTATTTTGTGCTCTCTACCTACCTATTAACACATGCTGCTTGTGGGTTTGTGTTTGT 3120
QY 3121 ATGTGCGTGTATACAGTATGTTAAATTTCCATGCGAGAAATTAATGCTGCTGATTTCT 3180
Db 3121 ATGTGCGTGTATACAGTATGTTAAATTTCCATGCGAGAAATTAATGCTGCTGATTTCT 3180
QY 3181 CAAAAAAGGCTGTTTGTGCTTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3206
Db 3181 CAAAAAAGGCTGTTTGTGCTTAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3206

RESULT 4

ADL62788
ID ADL62788 standard; DNA; 3499 BP.
XX
AC ADL62788;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human ovarian cancer DNA marker #21000.
XX
KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
OS Homo sapiens.
XX
PN WO200170979-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009126.

XX 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
PI Lee J, Lillie J;
XX
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
PS Disclosure; SEQ ID NO 21000; 106pp; English.
XX
CC The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
SQ Sequence 3499 BP; 1104 A; 610 C; 736 G; 1039 T; 0 U; 10 Other;
Query Match 98.8%; Score 3166; DB 5; Length 3499;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3169; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 11 GCGTCGGCGGAGCGACGCGCGCGCGCGCGCGGTACCTCTGCGCGCGGTCTCGCTCTC 70
DB 87 GAGCGCTGGCGAGCGACGCGCGCGCGCGCGGTACCTCTGCGCGCGGTCTCGCTCTC 146
QY 71 GCGCGGGCGGCGCGCGCGCGCGCGCGGTACCGCGCGCGCGCGCGCGGTACCTATATAA 130
DB 147 GCGCGGGCGGCGCGCGCGCGCGCGCGGTACCGCGCGCGCGCGCGCGGTACCTATATAA 206
QY 131 ATAAATCCATCAGAAATGACACCTTCTCAGGTGCGCTTGGAAATAGAGAACTCTTTTAC 190
DB 207 ATAAATCCATCAGAAATGACACCTTCTCAGGTGCGCTTGGAAATAGAGAACTCTTTTAC 266
QY 191 CAGGAGAAAGTTTTTGGCGATATGTGGAAAGCTGTGATGCTTTGGGAAACTGGAACTCTTCAA 250
DB 267 CAGGAGAAAGTTTTTGGCGATATGTGGAAAGCTGTGATGCTTTGGGAAACTGGAACTCTTCAA 326

QY 251 ATGCTGTGGCTCTTCTCCAGAGAAATGACACAGGTGAAAGCATGCTATGGAAAGCAACCA 310
DB 327 ATGCTGTGGCTCTTCTCCAGAGAAATGACACAGGTGAAAGCATGCTATGGAAAGCAACCA 386
QY 311 TTGTACTCAGTAGAGAGGATATCAGTTTCAGTATCGCTACTTTCAAAGGGTACTTTTGAAC 370
DB 387 TTGTACTCAGTAGAGAGGATATCAGTTTCAGTATCGCTACTTTCAAAGGGTACTTTTGAAC 446
QY 371 CAAAGACTATCGGTGGTCCATGTCAGTGATAGTTTCAAAAGTGGGAGACTCATCTACAC 430
DB 447 CAAAGACTATCGGTGGTCCATGTCAGTGATAGTTTCAAAAGTGGGAGACTCATCTACAC 506
QY 431 CAGGATCAATAAACCCTTTAGAAAGCGAAATTTATTGACCATGGAACAATTTGGAATCC 490
DB 507 CAGGATCAATAAACCCTTTAGAAAGCGAAATTTATTGACCATGGAACAATTTGGAATCC 566
QY 491 ACAATGGTGTGAAACTCTGGAATTTGGAATGGCTGACATGTGAGACTGAAATAAGATTAC 550
DB 567 ACAATGGTGTGAAACTCTGGAATTTGGAATGGCTGACATGTGAGACTGAAATAAGATTAC 626
QY 551 GTTTGCATTATTCTGAAAAACCTCTGTGTCAATTAACCAAGAAATAATTAATAATCTA 610
DB 627 GTTTGCATTATTCTGAAAAACCTCTGTGTCAATTAACCAAGAAATAATTAATAATCTA 686
QY 611 GATTAGGTGAGCTGACATAGAGGCTCTGAGGAAGATGACGATGATAGGATATCTC 670
DB 687 GATTAGGTGAGCTGACATAGAGGCTCTGAGGAAGATGACGATGATAGGATATCTC 746
QY 671 CCACTGTACTCCACAAAATGTCCAATAGCTTGGAGATATCCTTTAATAGCGCAATGAGT 730
DB 747 CCACTGTACTCCACAAAATGTCCAATAGCTTGGAGATATCCTTTAATAGCGCAATGAGT 806
QY 731 TCAAGTGCAGGCAATCACAGCCGAGTGTGTTATGCTTGCAGCTGATCGTTGACAG 790
DB 807 TCAAGTGCAGGCAATCACAGCCGAGTGTGTTATGCTTGCAGCTGATCGTTGACAG 866
QY 791 AGTACAGCATACAGAGATGGAACCACTGGAACCTAACTTTTGTGTTTTCGAAG 850
DB 867 AGTACAGCATACAGAGATGGAACCACTGGAACCTAACTTTTGTGTTTTCGAAG 926
QY 851 AAGATCTCAGTGAGCAGCTAGTTTCAAGGTGATGCCCTTCTGGACATGTGGGTACAGCTT 910
DB 927 AAGATCTCAGTGAGCAGCTAGTTTCAAGGTGATGCCCTTCTGGACATGTGGGTACAGCTT 986
QY 911 GTCTCTTATCATCCACATTTGCTGAGAGTGGAAAGAGTGTGAAATTTCTTACTCTTCCCA 970
DB 987 GTCTCTTATCATCCACATTTGCTGAGAGTGGAAAGAGTGTGAAATTTCTTACTCTTCCCA 1046
QY 971 TCATGACAGAAATTTCCGGAACCAATAGGAAAGTGTGAGTTGACTATATAATTTATTA 1030
DB 1047 TCATGACAGAAATTTCCGGAACCAATAGGAAAGTGTGAGTTGACTATATAATTTATTA 1106
QY 1031 AGCCATTACAGGATACAGTTGTGACATGAAATCTTTCATTTTCCAAGTATTGGAAGCCAA 1090
DB 1107 AGCCATTACAGGATACAGTTGTGACATGAAATCTTTCATTTCCAAGTATTGGAAGCCAA 1166
QY 1091 GAATACCATTTGGATTTGGCCATCGAGTGCAGGAAACTCTCAAACTGCCAGCTGG 1150
DB 1167 GAATACCATTTGGATTTGGCCATCGAGTGCAGGAAACTCTCAAACTGCCAGCTGG 1226
QY 1151 CTAAGTTCAAGAAATACTATTGCTTTTAAAGAAATGCTGTAGTTCATGTCAGCCT 1210
DB 1227 CTAAGTTCAAGAAATACTATTGCTTTTAAAGAAATGCTGTAGTTCATGTCAGCCT 1286
QY 1211 TTGTAGAAATTTGAGTACACCTTTCAAAGGACTTTGTGCCGTGATATCATGATCTTA 1270
DB 1287 TTGTAGAAATTTGAGTACACCTTTCAAAGGACTTTGTGCCGTGATATCATGATCTTA 1346
QY 1271 CCTGTGTTTGCATATGAAAAAGAAATTTGATGCTGATCGAGTTGAAATTTTGAATTC 1330
DB 1347 CCTGTGTTTGCATATGAAAAAGAAATTTGATGCTGATCGAGTTGAAATTTTGAATTC 1406

1331 CAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCAC 1390
1407 CAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCAC 1466
1391 TGAATCTTAAGGATCGGAAGAATCTGTGTTCCAGGAGGAATTCCTTTTCAGAAATC 1450
1467 TGAATCTTAAGGATCGGAAGAATCTGTGTTCCAGGAGGAATTCCTTTTCAGAAATC 1526
1451 AGCCATTTCCCTCTCTTAAGATGGTTTATAGATCTTTGCGAGGAAGATGTAGGGTTTAA 1510
1527 AGCCATTTCCCTCTCTTAAGATGGTTTATAGATCTTTGCGAGGAAGATGTAGGGTTTAA 1586
1511 TTGAATAAATGGATCTGCCAGCAAAAGGATGGAATGTGGGATGTAACATTTATCAACAT 1570
1587 TTGAATAAATGGATCTGCCAGCAAAAGGATGGAATGTGGGATGTAACATTTATCAACAT 1646
1571 ATTTTCACATGAATCTGTTTTCGATATATTTTAAAGCTGTTTATAGAAATCTCGGCA 1630
1647 ATTTTCACATGAATCTGTTTTCGATATATTTTAAAGCTGTTTATAGAAATCTCGGCA 1706
1631 AGAGGAGAATAGTGTGTTTCTTCAATTTGATGCGAGATATTTGCACAATGTTTCGCAAAAGC 1690
1707 AGAGGAGAATAGTGTGTTTCTTCAATTTGATGCGAGATATTTGCACAATGTTTCGCAAAAGC 1766
1691 AGAACAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGGAATCA 1750
1767 AGAACAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGGAATCA 1826
1751 TGGACCTCAGATCTCGGCAACCCCAATTCGAATGAGCTTCGACAGTTTGAATCTAC 1810
1827 TGGACCTCAGATCTCGGCAACCCCAATTCGAATGAGCTTCGACAGTTTGAATCTAC 1886
1811 TGGGATAAATGTACATCTGAAGACTGTCTCAGAAACCCATCCTATATTTCAAGAGGCAA 1870
1887 TGGGATAAATGTACATCTGAAGACTGTCTCAGAAACCCATCCTATATTTCAAGAGGCAA 1946
1871 AGCTAAGGACTAGTATATTTCTGCTGGGGTATGATACCAATGATCTGGAAGAGAA 1930
1947 AAGCTAAGGACTAGTATATTTCTGCTGGGGTATGATACCAATGATCTGGAAGAGAA 2006
1931 GGAATTTGAAGAACTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATGGATGC 1990
2007 GGAATTTGAAGAACTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATGGATGC 2066
1991 CTGAACAAACCAATATATTTCCAGTGGAGCAATTTGGAAGCCCTGAAGCAGGAATGCCAG 2050
2067 CTGAACAAACCAATATATTTCCAGTGGAGCAATTTGGAAGCCCTGAAGCAGGAATGCCAG 2126
2051 AGCTTAAGAGCTGTTTGTGTCACATGTTAGCGGCTTGTTCCTCATCTTTGTGGGG 2110
2127 AGCTTAAGAGCTGTTTGTGTCACATGTTAGCGGCTTGTTCCTCATCTTTGTGGGG 2186
2111 AGTCTGATATCCATGTGGATGCCAACGGCAATTTGATACCTGGAGATGCTTAGTTTAT 2170
2187 AGTCTGATATCCATGTGGATGCCAACGGCAATTTGATACCTGGAGATGCTTAGTTTAT 2246
2171 TGCAACAGAGTCAATTTTGGGGGCTGCACCGCTGTTCTGGGTATTCATTTTTCATCACTG 2230
2247 TGCAACAGAGTCAATTTTGGGGGCTGCACCGCTGTTCTGGGTATTCATTTTTCATCACTG 2306
2231 AGCATTTGATCTATGCTCTTTGGGCTTCTCAGTTCAATGAAGCAATTAATGAAGTATTT 2290
2307 AGCATTTGATCTATGCTCTTTGGGCTTCTCAGTTCAATGAAGCAATTAATGAAGTATTT 2366
2291 AACTCTTTCACATCACTGTTTGAAGTATGCTATTTTAAATTAATCTTGGCCAGGATTAATG 2350
2367 AACTCTTTCACATCACTGTTTGAAGTATGCTATTTTAAATTAATCTTGGCCAGGATTAATG 2426
2351 CCAGTCAGTCTCTTTATAGTGAGAAATTTTATTTGGTTAGTATATAAATTTTAAACTA 2410
2427 CCAGTCAGTCTCTTTATAGTGAGAAATTTTATTTGGTTAGTATATAAATTTTAAACTA 2486
2411 AATATATAAATCTATATGTTTAAACATATGTTTCAATTTAAAGCATAGCACTTTGAAATTA 2470

2487 AATATATAAATCTATATGTTTAAACATATGTTTCAATTTAAAGCATAGCACTTTGAAATTA 2546
2471 CTATATAAATAGCTCATATTTTACACTTACAGCTTTTCACTTGTGATCAGGTCTGAAATCTTT 2530
2547 CTATATAAATAGCTCATATTTTACACTTACAGCTTTTCACTTGTGATCAGGTCTGAAATCTTT 2606
2531 AGCACTTAAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2590
2607 AGCACTTAAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2666
2591 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2650
2667 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2726
2651 AGGATATTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2710
2727 AGGATATTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2786
2711 AGGCTGCTGCCATTTTCCACTTAAACCAACAGCTGAAAGTGAACCTCGAAATCTGTT 2770
2787 AGGCTGCTGCCATTTTCCACTTAAACCAACAGCTGAAAGTGAACCTCGAAATCTGTT 2846
2771 TCATATAATCTTCAAAAGTGTGTTTACATCAATGTTTAAATTTCAAAATGCTGCAAGGTA 2830
2847 TCATATAATCTTCAAAAGTGTGTTTACATCAATGTTTAAATTTCAAAATGCTGCAAGGTA 2906
2831 ATTTAATGTATAAATATTTAGTAAAGAAAGTATGTTTGCATCTTGTAGTAGAATAGATC 2890
2907 ATTTAATGTATAAATATTTAGTAAAGAAAGTATGTTTGCATCTTGTAGTAGAATAGATC 2966
2891 ACAACATCAAAATTCATTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2950
2967 ACAACATCAAAATTCATTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3026
2951 TACTGTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3010
3027 TACTGTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3086
3011 CGTATTTTGTGCTTAAAGAAAGGCTGTTTGTAGGCTTTTAAATATGCTTATTTT 3070
3087 CGTATTTTGTGCTTAAAGAAAGGCTGTTTGTAGGCTTTTAAATATGCTTATTTT 3146
3071 TGTGCTCTCACTACTTATACACACTGTTGCTTTTGTGGGTTTGTGTTGATGCTGCTGT 3130
3147 TGTGCTCTCACTACTTATACACACTGTTGCTTTTGTGGGTTTGTGTTGATGCTGCTGT 3206
3131 GTTATACAGTGTAAATTTTCCATGCAAGAAATTAATGCTGCTGAAATCTCTCAA 3184
3207 GTTATACAGTGTAAATTTTCCATGCAAGAAATTAATGCTGCTGAAATCTCTCAA 3260

RESULT 5

ADN05171

ID ADN05171 standard; cDNA; 5443 BP.

XX AC ADN05171;

XX AC ADN05171;

DT 01-JUL-2004 (first entry)

DE Antipsoriatic cDNA sequence #801.

XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

XX Homo sapiens.

XX WO2004028479-A2.

XX PD 08-APR-2004.

XX 25-SEP-2003; 2003WO-US030907.

XX 25-SEP-2002; 2002US-0414006P.

1811 TGGGGATAAATGTATCATCTGAAGACTTCTCTCAGAAAACCATCCCTATATTCAGAAGGCAA 1870
1871 TGGGGATAAATGTATCATCTGAAGACTTCTCTCAGAAAACCATCCCTATATTCAGAAGGCAA 1930
1871 AAGCTAAGGGACTAGTCATATTCCTGCTGGGTGATGATACCAATGATCCTGAACACAGAA 1930
1931 AAGCTAAGGGACTAGTCATATTCCTGCTGGGTGATGATACCAATGATCCTGAACACAGAA 1990
1931 GGAATTTGAAGGAACCTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATGATGC 1990
1991 GGAATTTGAAGGAACCTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATGATGC 2050
1991 CTGAACAACCAATATATTTCAAGTGGAGCAATTTGAAGCCCTGAAGCAAGAAATGGCCAG 2050
2051 CTGAACAACCAATATATTTCAAGTGGAGCAATTTGAAGCCCTGAAGCAAGAAATGGCCAG 2110
2051 AGCTTAAGAGCTGTTTGTGTCCTCCACCTGTTAGCGCTTGTTCCTCCATCTTTGTGGGG 2110
2111 AGCTTAAGAGCTGTTTGTGTCCTCCACCTGTTAGCGCTTGTTCCTCCATCTTTGTGGGG 2170
2111 AGCTGATATCCATGTGATGTCACACCGCATTTGATAACCTGGAGATGCTTAGTTTAT 2170
2171 AGCTGATATCCATGTGATGTCACACCGCATTTGATAACCTGGAGATGCTTAGTTTAT 2230
2171 TGCACAGAGGTCAATTTTGGGGCGGTGCACCGCTGTTCTGGGTATTCAATTTTCACTACTG 2230
2231 TGCACAGAGGTCAATTTTGGGGCGGTGCACCGCTGTTCTGGGTATTCAATTTTCACTACTG 2290
2231 AGCATTTGTATCTATGCCCTTTTGGGCTTCTCAGTTCATGAAGCAATGAATGATTT 2290
2291 AGCATTTGTATCTATGCCCTTTTGGGCTTCTCAGTTCATGAAGCAATGAATGATTT 2350
2291 AACTCTTTTCACTACAGTCTCTGCAAGTATGCTATTTAAATTAATCTGGCCAGGTATAATG 2350
2351 AACTCTTTTCACTACAGTCTCTGCAAGTATGCTATTTAAATTAATCTGGCCAGGTATAATG 2410
2351 CCAGTCAGTCTCTTTATAGTGAGAAAATTTATTTGGTTAGTAAATATAATTTTAAACTA 2410
2411 CCAGTCAGTCTCTTTATAGTGAGAAAATTTATTTGGTTAGTAAATATAATTTTAAACTA 2470
2411 AATATATAAATCTATAATTTAAACATATGTTTCATTTAAAGCATAGCACCTTTGAAATTA 2470
2471 AATATATAAATCTATAATTTAAACATATGTTTCATTTAAAGCATAGCACCTTTGAAATTA 2530
2471 CTATATAAATAGTCTCATATTTTACATTTACAGTCTTTTCAATTTGATCAGGTCTGAAATCTTT 2530
2531 CTATATAAATAGTCTCATATTTTACATTTACAGTCTTTTCAATTTGATCAGGTCTGAAATCTTT 2590
2531 AGCATTAAAGGAAAATGACTATGCATAATTTATACCTGACCATGAAAAAATAAGTACCTC 2590
2591 AGCATTAAAGGAAAATGACTATGCATAATTTATACCTGACCATGAAAAAATAAGTACCTC 2650
2591 AAATGATGATTTGACATGCTGATTTCCAACTGCAAAATCTTTGTCATCTTGTATAT 2650
2651 AAATGATGATTTGACATGCTGATTTCCAACTGCAAAATCTTTGTCATCTTGTATAT 2710
2651 AGGTATTTTACATGCTGATTTGACATGCTGATTTCCAACTGCAAAATCTTTGTCATCTTGTATAT 2710
2711 AGGTATTTTACATGCTGATTTGACATGCTGATTTCCAACTGCAAAATCTTTGTCATCTTGTATAT 2770
2711 AGGCTGCTGCCATTTTCCACTTAACCAACACAGCTGAAAGTGAACCTCGAAATCTGTT 2770
2771 AGGCTGCTGCCATTTTCCACTTAACCAACACAGCTGAAAGTGAACCTCGAAATCTGTT 2830
2771 TCATAAATCTTTCAAAAGTTGTTTACATCAATGTTAAATTTCAAAATCTGCGAGGTA 2830
2831 TCATAAATCTTTCAAAAGTTGTTTACATCAATGTTAAATTTCAAAATCTGCGAGGTA 2890
2831 ATTTAATGATATAAATATTTAGTAAAGAAAGTATGTTGATGCTACTTGTAGTAAATAGATC 2890
2891 ATTTAATGATATAAATATTTAGTAAAGAAAGTATGTTGATGCTACTTGTAGTAAATAGATC 2950

2891 ACAACATACAAATTCAAATTCAGTGCATGCTTTAGGTGTAAAGCATGAGATTGTACATGTT 2950
2951 ACAACATACAAATTCAAATTCAGTGCATGCTTTAGGTGTAAAGCATGAGATTGTACATGTT 3010
2951 TACTGTTAGGTCCCTTGGCATCTGCTGCTAGTGTAGTATGAGAAGCATGTCAGAGGACTGGA 3010
3011 TACTGTTAGGTCCCTTGGCATCTGCTGCTAGTGTAGTATGAGAAGCATGTCAGAGGACTGGA 3070
3011 CGTATTTTGTTCCTTAAAAAAAAGGCTGTTTGTAGCGTTTTAAATAATGCTTATTTTG 3070
3071 CGTATTTTGTTCCTTAAAAAAAAGGCTGTTTGTAGCGTTTTAAATAATGCTTATTTTG 3130
3071 TGTGCTCTCCTACTACTATACACACTGCTTGTGCTTGTGGGTTTGTGTATGTCGGTGT 3130
3131 TGTGCTCTCCTACTACTATACACACTGCTTGTGCTTGTGGGTTTGTGTATGTCGGTGT 3190
3131 GTTATACAGTAGTTAAATTTCCATGCAGAAAAATAAATGCTCTGAATTTCTCAA 3184
3191 GTTATACAGTAGTTAAATTTCCATGCAGAAAAATAAATGCTCTGAATTTCTCAT 3244

RESULT 6
ADS74320
ID ADS74320 standard; cDNA; 5443 BP.
XX
AC ADS74320;
XX AC
XX DT 16-DEC-2004 (first entry)
XX
DE PRO83903 cDNA clone DNA327983, role in immune-related disease.
XX
KW PRO83903; Crohn's disease; rheumatoid arthritis; gastrointestinal-gen.;
XX antirheumatic; antiarthritic; psoriasis; antipsoriatic; gene; ss.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 190..2223
FT /*tag= a
FT /product= "PRO83903"
FT /partial
FT /note= "No start codon"
XX
PN WO2004081199-A2.
XX
PD 23-SEP-2004.
XX
PF 10-MAR-2004; 2004WO-US007862.
XX
PR 11-MAR-2003; 2003US-0454025P.
XX (GETH) GENENTECH INC.
XX
PI Baldwin D, Bodary S, Clark H, Pong S, Gurney AL, Williams PM;
XX
XX WPI; 2004-668955/65.
XX P-PSDB; ADS74321.
XX
PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating psoriasis, Crohn's disease, Ulcerative Colitis, or rheumatoid
PT arthritis.
XX
PS Claim 1; SEQ ID NO 39; 166pp; English.
XX
CC The present sequence is of cDNA clone DNA327983 encoding novel human PRO
CC polypeptide. PRO83903. The invention provides newly identified and
CC isolated nucleotide sequences encoding polypeptides referred to as PRO
CC polypeptides that are useful in the diagnosis and treatment of immune-
CC related diseases. Microarray analysis showed that DNA327983 is up-
CC regulated 1.5-fold in lesional skin as compared to non-lesional skin from
CC psoriasis patients, up-regulated 1.3-fold in colon samples from Crohn's
CC disease patients as compared to normal colon and up-regulated 1.3-fold in
CC white blood cells from rheumatoid arthritis patients as compared to those

from healthy donors. It is also down-regulated 1.5-fold upon activation of CD4 T cells with CD28 or ICAM and down-regulated 2-fold upon activation of monocytes with LPS. PR083903 can be used in a claimed method of identifying a compound that inhibits expression of the gene encoding it. The candidate compound is especially an antisense nucleic acid. The PR0 polypeptide, its antagonist or an antibody that binds the polypeptide are used in claimed methods for the alleviation or diagnosis of rheumatoid arthritis, Crohn's disease and psoriasis. A vector comprising the present nucleic acid can be used to transform a host cell, especially a CHO cell, *Escherichia coli* or yeast, for production of the PR0 polypeptide.

XX	Sequence	5443 BP; 1544 A; 1034 C; 1149 G; 1716 T; 0 U; 0 Other;
SQ	Query Match	98.8%; Score 3166; DB 13; Length 5443;
	Best Local Similarity	99.8%; Pred. No. 0;
	Matches 3169; Conservative	0; Mismatches 5; Indels 0; Gaps 0;
QY	11	GCCTCCGGGCGAGGACGGACGGCGGGGCCCGGTACCTCTGCCCGGGTCTCGCTCTC 70
DB	71	GAGCGCTGGCGAGGACGGACGGCGGGGCCCGGTACCTCTGCCCGGGTCTCGCTCTC 130
QY	71	GGGCGGGCGGGCGAGCGCGGACCTGCGGACTAGCGAACCCGGAGACGACATCATAAA 130
DB	131	GGGCGGGCGGGCGAGCGCGGACCTGCGGACTAGCGAACCCGGAGACGACATCATAAA 190
QY	131	ATAAATCCATCAGAAATGACACCTCTCAGGTTGCCCTTTGAAATAAGAGGAACCTCTTTTAC 190
DB	191	ATAAATCCATCAGAAATGACACCTCTCAGGTTGCCCTTTGAAATAAGAGGAACCTCTTTTAC 250
QY	191	CAGGAGAAAGTTTTTTCGGATATGTGGAAGCTGTGATGCTTTTGGGAAACTGGGAATCTCTAAA 250
DB	251	CAGGAGAAAGTTTTTTCGGATATGTGGAAGCTGTGATGCTTTTGGGAAACTGGGAATCTCTAAA 310
QY	251	ATGCTGTGGCTCTTCTTCCAGAAATGACACAGGTGAAAGCATGCTATGGAAGCAACCA 310
DB	311	ATGCTGTGGCTCTTCTTCCAGAAATGACACAGGTGAAAGCATGCTATGGAAGCAACCA 370
QY	311	TTGTACTCAGTAGAGGAGTATCAGTTCAGTATCGCTACTTTCAAAGGGTACTTTTTAGAAC 370
DB	371	TTGTACTCAGTAGAGGAGTATCAGTTCAGTATCGCTACTTTCAAAGGGTACTTTTTAGAAC 430
QY	371	CAAAGACTATCCGCTGGTCCATGCTCAAGTGATATAGTTTCAAGTGGGAGACTCATCTACAAC 430
DB	431	CAAAGACTATCCGCTGGTCCATGCTCAAGTGATATAGTTTCAAGTGGGAGACTCATCTACAAC 490
QY	431	CA CGA TCAATAAACCCCTTTAGAAAAGCGAAATTTATTTA TGA CGATGGACAAATTTGGAATCC 490
DB	491	CACGA TCAATAAACCCCTTTAGAAAAGCGAAATTTATTTA TGA CAGATGGACAAATTTGGAATCC 550
QY	491	ACAATGGTGTGAAAACCTCTGGATTCCTGATGCGTACATGTCAGACTGGAATAAGATTAC 550
DB	551	ACAATGGTGTGAAAACCTCTGGATTCCTGATGCGTACATGTCAGACTGGAATAAGATTAC 610
QY	551	GTTTTGCATTTATCTGAAAAACCTCCTGTGTCAATAACCAAGAAAAAATTTAAAAAATCTTA 610
DB	611	GTTTTGCATTTATCTGAAAAACCTCCTGTGTCAATAACCAAGAAAAAATTTAAAAAATCTTA 670
QY	611	GATTTAGGGTGAAGCTGACACTAGAGGGCTGGAGGAAGATGACGATGATAGGTTATCTC 670
DB	671	GATTTAGGGTGAAGCTGACACTAGAGGGCTGGAGGAAGATGACGATGATAGGTTATCTC 730
QY	671	CCACTGTATCTCGCAAAAATGTCCAATAGCTTTGAGATATCCTTAATAAGCGCAAAATGAGT 730
DB	731	CCACTGTATCTCGCAAAAATGTCCAATAGCTTTGAGATATCCTTAATAAGCGCAAAATGAGT 790
QY	731	TCAAGTGCAGGCATTCACACGCCGAGTGTGGTTATGGCTTGACGCTGATCGTTGGACAG 790
DB	791	TCAAGTGCAGGCATTCACACGCCGAGTGTGGTTATGGCTTGACGCTGATCGTTGGACAG 850
QY	791	AGTACAGCATACAGACGATGGAAACCGAATAACCTGGAACTAAATCTTTGATTTTTTTCGAAG 850
DB	851	AGTACAGCATACAGACGATGGAAACCGAATAACCTGGAACTAAATCTTTGATTTTTTTCGAAG 910

QY	851	AAGATCTCAGTGA	GACCACTAGTTC	CAGGCTGA	TGCCCTTC	TGGA	CA	TGTGGGTA	CAGCTT	910
DB	911	AAATCTCAGTGA	GACCACTAGTTC	CAGGCTGA	TGCCCTTC	TGGA	CA	TGTGGGTA	CAGCTT	970
QY	911	GTCTCTTATCAT	CCACATTC	GTGAGAGT	GGAAGAGT	GTGGAA	TTCTT	ACTCTTCCCA		970
DB	971	GTCTCTTATCAT	CCACATTC	GTGAGAGT	GGAAGAGT	GTGGAA	TTCTT	ACTCTTCCCA		1030
QY	971	TCATGACAGAAA	TTCCCGGAAA	CAATAGC	CAAGTGCAGAT	TATATA	TAATTA	TTA		1030
DB	1031	TCATGACAGAAA	TTCCCGGAAA	CAATAGC	CAAGTGCAGAT	TATATA	TAATTA	TTA		1090
QY	1031	AGCCATTACCA	GATACAGTTC	GTGATGAA	TCTTCA	TTTCCAA	GTATTTGGA	AGCCAA		1090
DB	1091	AGCCATTACCA	GATACAGTTC	GTGATGAA	TCTTCA	TTTCCAA	GTATTTGGA	AGCCAA		1150
QY	1091	GAATACCA	TTGGATGTGGCCAT	TCAGAGTGC	AGGAAA	CTCTACA	CAACTGCC	AGCTGG		1150
DB	1151	GAATACCA	TTGGATGTGGCCAT	TCAGAGTGC	AGGAAA	CTCTACA	CAACTGCC	AGCTGG		1210
QY	1151	CTAAAGTTC	CAAGAAATACTA	TTGCTCTTTA	AGAAATGCT	CTAGTCA	TGTG	CAGCT		1210
DB	1211	CTAAAGTTC	CAAGAAATACTA	TTGCTCTTTA	AGAAATGCT	CTAGTCA	TGTG	CAGCT		1270
QY	1211	TTGTGAA	TTTGACGTAC	CACTTTCAA	AGGACTTTGTG	CCGTGGTATA	TCATGATCTTA			1270
DB	1271	TTGTGAA	TTTGACGTAC	CACTTTCAA	AGGACTTTGTG	CCGTGGTATA	TCATGATCTTA			1330
QY	1271	CTGTGTTGTT	TGACTATG	AAAAAGAA	ATTGATGCTG	ATCCAGT	TGAA	TTTGTGAA	ATTC	1330
DB	1331	CTGTGTTGTT	TGACTATG	AAAAAGAA	ATTGATGCTG	ATCCAGT	TGAA	TTTGTGAA	ATTC	1390
QY	1331	CAGTAA	AGAAATTA	ACATTTGACCA	ACTCCAGT	GTGTTAA	AGCTCACT	CATGTG	ACTGAC	1390
DB	1391	CAGTAA	AGAAATTA	ACATTTGACCA	ACTCCAGT	GTGTTAA	AGCTCACT	CATGTG	ACTGAC	1450
QY	1391	TGAAATCTA	AGGATCGGA	AGAAATCTGTG	TTCCAGG	AGGAAATTCCT	TTTTCAG	AAAAATC		1450
DB	1451	TGAAATCTA	AGGATCGGA	AGAAATCTGTG	TTCCAGG	AGGAAATTCCT	TTTTCAG	AAAAATC		1510
QY	1451	AGCCATTC	CTCTCTT	TAAGATGG	TTTAGAGTCT	TGTCAG	AGAGATGTAG	GGTTTAA	CA	1510
DB	1511	AGCCATTC	CTCTCTT	TAAGATGG	TTTAGAGTCT	TGTCAG	AGAGATGTAG	GGTTTAA	CA	1570
QY	1511	TTGAAAT	AAAAATG	GAATCTGCC	CAAGGATGGA	ATGTGG	ATGTA	CTTATCA	CA	1570
DB	1571	TTGAAAT	AAAAATG	GAATCTGCC	CAAGGATGGA	ATGTGG	ATGTA	CTTATCA	CA	1630
QY	1571	ATTTTGAT	CA	TGAATCTG	TTTTTGATATA	TTTTTAAAA	ACTGTTTT	GAGAAA	TTCTGGGA	1630
DB	1631	ATTTTGAT	CA	TGAATCTG	TTTTTGATATA	TTTTTAAAA	ACTGTTTT	GAGAAA	TTCTGGGA	1690
QY	1631	AGAGG	AGAAATG	TGTTTTCTT	CAATTC	GATGAGAT	TTTGCA	CAATGG	TTTCGGCA	1690
DB	1691	AGAGG	AGAAATG	TGTTTTCTT	CAATTC	GATGAGAT	TTTGCA	CAATGG	TTTCGGCA	1750
QY	1691	AGAA	CAAAAT	TCGAT	ACTATTTTAA	CTCAAGG	AAATCTG	GATTTAT	CTCGAA	1750
DB	1751	AGAA	CAAAAT	TCGAT	ACTATTTTAA	CTCAAGG	AAATCTG	GATTTAT	CTCGAA	1810
QY	1751	TGGAC	CTCAGAT	CTCGG	CAACCCCA	TTGCAAT	TGAC	TTTGA	CTTCTG	1810
DB	1811	TGGAC	CTCAGAT	CTCGG	CAACCCCA	TTGCAAT	TGAC	TTTGA	CTTCTG	1870
QY	1811	TGGGG	TAAAT	TGTA	CTGA	AGCTTG	CTCAG	AAACCA	TTCTAT	1870
DB	1871	TGGGG	TAAAT	TGTA	CTGA	AGCTTG	CTCAG	AAACCA	TTCTAT	1930
QY	1871	AAGCT	TAAGG	GA	CTAGT	CA	TTCTG	CTGGG	TGAT	1930
DB	1931	AAGCT	TAAGG	GA	CTAGT	CA	TTCTG	CTGGG	TGAT	1990

Db 87 GAGCGCTGGCGAGGCA CGGACGGCGGGCGCCGGTACCTCTGCCCCGGTCTCGCTCTC 146
Qy 71 GGGCGGGCGCGCGGAGCGCGGACCTGCGGACTAGCGAA CCGGAGGACGACATCATAAA 130
Db 147 GGGCGGGCGCGCGGAGCGCGGACCTGCGGACTAGCGAA CCGGAGGACGACATCATAAA 206
Qy 131 ATAAATCCATCAGAAATGACACTTCTCAGGTTGCTTGTGAATTAAGAGGAACCTCTTTTAC 190
Db 207 ATAAATCCATCAGAAATGACACTTCTCAGGTTGCTTGTGAATTAAGAGGAACCTCTTTTAC 266
Qy 191 CAGGAGAAATTTTGGCATATGTGGAGCTGTGATGCTTTTGGGAAA CTGGAATCCTCAAA 250
Db 267 CAGGAGAAATTTTGGCATATGTGGAGCTGTGATGCTTTGGA AAA CTGGAA TCCCTCAA 326
Qy 251 ATGCTGTGGCTCTTCTCCAGAGAAATGACACAGGTGAAGCATGCTATGGAAGCAACCA 310
Db 327 ATGCTGTGGCTCTTCTCCAGAGAAATGACACAGGTGAAGCATGCTATGGAAGCAACCA 386
Qy 311 TTGTACTCAGTAGAGGATATCAGTTTCAATGCTGCTACTTCAAAAGGGTACTTTTTAGAAC 370
Db 387 TTGTACTCAGTAGAGGATATCAGTTTCAATGCTGCTACTTCAAAAGGGTACTTTTTAGAAC 446
Qy 371 CAAGACTATCGGTGCTCCATGTCAAGTGTAGTTTCAAGTGGGAGACTCATCTACAAC 430
Db 447 CAAGACTATCGGTGCTCCATGTCAAGTGTAGTTTCAAGTGGGAGACTCATCTACAAC 506
Qy 431 CACGATCAATAACCCCTTTAGAAAGCGAAATTTATTTAGCAGTGGACAAATTTGGAATCC 490
Db 507 CACGATCAATAACCCCTTTAGAAAGCGAAATTTATTTAGCAGTGGACAAATTTGGAATCC 566
Qy 491 ACAATGGTGTGAAACTCTGGATTTCTGGATGCTGACATGTTCAGACTGGAATAGATTAC 550
Db 567 ACAATGGTGTGAAACTCTGGATTTCTGGATGCTGACATGTTCAGACTGGAATAGATTAC 626
Qy 551 GTTTGCAATTTCTGAAAAACCTCTGCTCAATAACCAAGAAATAATTTAAAAAATCTA 610
Db 627 GTTTGCAATTTCTGAAAAACCTCTGCTCAATAACCAAGAAATAATTTAAAAAATCTA 686
Qy 611 GATTTAGGCTGAAGCTGACACTAGAAAGCCTGGAGGAAAGATGACGATAGAGGTATCTC 670
Db 687 GATTTAGGCTGAAGCTGACACTAGAAAGCCTGGAGGAAAGATGACGATAGAGGTATCTC 746
Qy 671 CCACGTACTCCAAAAATGTCAAATAGCTTGGAGATATCCTTAATAAGCACAATGAT 730
Db 747 CCACGTACTCCAAAAATGTCAAATAGCTTGGAGATATCCTTAATAAGCACAATGAT 806
Qy 731 TCAAGTGCAGGCATTTCAAGCCGAGTGTGTTATGGCTTGAGCTGATCGTTGGACAG 790
Db 807 TCAAGTGCAGGCATTTCAAGCCGAGTGTGTTATGGCTTGAGCTGATCGTTGGACAG 866
Qy 791 AGTACAGCATACAGACGATGGAAACAGATAAACCTGGAACTAATCTTTGATTTTTTCGAAG 850
Db 867 AGTACAGCATACAGACGATGGAAACAGATAAACCTGGAACTAATCTTTGATTTTTTCGAAG 926
Qy 851 AAGATCTCAGTAGGACGATGTTACAGGTGATGCCCCCTTCCTGGA CATGGGATAAGCTT 910
Db 927 AAGATCTCAGTAGGACGATGTTACAGGTGATGCCCCCTTCCTGGA CATGGGATAAGCTT 986
Qy 911 GTCTCTTATCATCCACCATTTCTGAGAGTGGAAAGAGTCTGGAATTTCTTACTCTTCCCA 970
Db 987 GTCTCTTATCATCCACCATTTCTGAGAGTGGAAAGAGTCTGGAATTTCTTACTCTTCCCA 1046
Qy 971 TCAAGACAGAAATTTCCGGAAAAAATAAGGCAAAAGTGAAGTTGACTATATAATTA 1030
Db 1047 TCAAGACAGAAATTTCCGGAAAAAATAAGGCAAAAGTGAAGTTGACTATATAATTA 1106
Qy 1031 AGCCATTTACCAGGATACAGTTGTGACATGAAATCTTTCAATTTTCCAAGTATTGGAAGCCAA 1090
Db 1107 AGCCATTTACCAGGATACAGTTGTGACATGAAATCTTTCAATTTTCCAAGTATTGGAAGCCAA 1166
Qy 1091 GAATACCATTTGGATTTGGCATTCGAGGTGCGAGGAACCTTACACAACTGCCAGCTGG 1150
Db 1167 GAATACCATTTGGATTTGGCATTCGAGGTGCGAGGAACCTTACACAACTGCCAGCTGG 1226

Qy 1151 CTAAAGTTCAAGAAAATACTATTGCTTCTTTAAGAAAATGCTCTAGTCAATGCTGAGCCT 1210
Db 1227 CTAAAGTTCAAGAAAATACTATTGCTTCTTTAAGAAAATGCTCTAGTCAATGCTGAGCCT 1286
Qy 1211 TTGTAGAAATTTGACGTFACACCTTTCAAAGGACTTTGTGCCCCGTGGTATATCATGATCTTA 1270
Db 1287 TTGTAGAAATTTGACGTFACACCTTTCAAAGGACTTTGTGCCCCGTGGTATATCATGATCTTA 1346
Qy 1271 CTTGTTGTTTGTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTGAAAATTC 1330
Db 1347 CTTGTTGTTTGTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTTGTGAAAATTC 1406
Qy 1331 CAGTAAAGAAATTAACATTTTGACCAACTCCAGTTGTTTAAAGCTCACTCATGTGACTGCAC 1390
Db 1407 CAGTAAAGAAATTAACATTTTGACCAACTCCAGTTGTTTAAAGCTCACTCATGTGACTGCAC 1466
Qy 1391 TGAATCTAAGGATCGGAAGAAATCTGTGTTTTCAGGAGGAAAAATTCCTTTTCAGAAAAATC 1450
Db 1467 TGAATCTAAGGATCGGAAGAAATCTGTGTTTTCAGGAGGAAAAATTCCTTTTCAGAAAAATC 1526
Qy 1451 AGCCATTTCTCTCTTAAGATGTTTAAAGTCTTTGCGCAGAGAAATGAGGGTTTAAACA 1510
Db 1527 AGCCATTTCTCTCTTAAGATGTTTAAAGTCTTTTGCAGAGAAATGAGGGTTTAAACA 1586
Qy 1511 TTGAATTAATAATGATCTGCGAGCAAGGATGGAATGTGGATGGAATCTTATCAACAT 1570
Db 1587 TTGAATTAATAATGATCTGCGAGCAAGGATGGAATGTGGATGGAATCTTATCAACAT 1646
Qy 1571 ATTTTGACATGAAATCTGTTTTTGGATATAATTTTAAAAAATCTGTTTTAGAAAAATCTGGGA 1630
Db 1647 ATTTTGACATGAAATCTGTTTTTGGATATAATTTTAAAAAATCTGTTTTAGAAAAATCTGGGA 1706
Qy 1631 AGAGGAAATGATGTTTTCTTCAATTTGATGCGAGATATTTGCACAATGGTTTCGCCAAAAGC 1690
Db 1707 AGAGGAAATGATGTTTTCTTCAATTTGATGCGAGATATTTGCACAATGGTTTCGCCAAAAGC 1766
Qy 1691 AGAACAAATATCCGACTACTATTTTAACTCAAGGAAAAATCTGAGATTTATCTCGAACTCA 1750
Db 1767 AGAACAAATATCCGACTACTATTTTAACTCAAGGAAAAATCTGAGATTTATCTCGAACTCA 1826
Qy 1751 TGGACCTCAGATCTCGGACAAACCCCAATTTGCAATGAGCTTTGCAAGTTTGAATAATCTTAC 1810
Db 1827 TGGACCTCAGATCTCGGACAAACCCCAATTTGCAATGAGCTTTGCAAGTTTGAATAATCTTAC 1886
Qy 1811 TGGGGATAAATGATACATGCTGAGAAAGCTGCTGAGAAACCCCATCTTATTCAGAGGCA 1870
Db 1887 TGGGGATAAATGATACATGCTGAGAAAGCTGCTGAGAAACCCCATCTTATTCAGAGGCA 1946
Qy 1871 AAGCTAAGGACTAGTCAATATCTGCTGGGGTGAATGATACCAATGATCCTGAAAAACAGAA 1930
Db 1947 AAGCTAAGGACTAGTCAATATCTGCTGGGGTGAATGATACCAATGATCCTGAAAAACAGAA 2006
Qy 1931 GGAATTTGAAGGAACTTTGGAGTTAAATGGTCTAAATTTATGATAGGATATATGATTTGATGC 1990
Db 2007 GGAATTTGAAGGAACTTTGGAGTTAAATGGTCTAAATTTATGATAGGATATATGATTTGATGC 2066
Qy 1991 CTGAACCAACCAATATATATTCAGAGTGGACAAATTTGAAACGCCCTGAAGCAGAAATTTGCCAG 2050
Db 2067 CTGAACCAACCAATATATATTCAGAGTGGACAAATTTGAAACGCCCTGAAGCAGAAATTTGCCAG 2126
Qy 2051 AGCTTAAAGACTGTTGTGTTGCCACTGTTAGCCGCTTTGTTCCCTCATCTTTGTGTGGGG 2110
Db 2127 AGCTTAAAGACTGTTGTGTTGCCACTGTTAGCCGCTTTGTTCCCTCATCTTTGTGTGGGG 2186
Qy 2111 AGTCTGATATCCATGTCGATGCCAAACCGCATTTGATAACGTGGGAATGCTTAGTTTTTAT 2170
Db 2187 AGTCTGATATCCATGTCGATGCCAAACCGCATTTGATAACGTGGGAATGCTTAGTTTTTAT 2246
Qy 2171 TGCAACAGAGGTCAATTTTGGGGGGGTGCAACCGCTGTTCTCGGGTATTCATTTTTTCACTG 2230
Db 2247 TGCAACAGAGGTCAATTTTGGGGGGGTGCAACCGCTGTTCTCGGGTATTCATTTTTTCACTG 2306

```
QY 2231 AGCATTTGATGATGCTTTTGGCTTCTCAGTTCAATGAAGCAATAATGAAGTATTT 2290
DB 2307 AGCATTTGATGATGCTTTTGGCTTCTCAGTTCAATGAAGCAATAATGAAGTATTT 2366
QY 2291 AACTCTTTTCACTACAGTCTTCTGCAAGTATGCTATTTAAATTAATTTGGCCAGGTATAATG 2350
DB 2367 AACTCTTTTCACTACAGTCTTCTGCAAGTATGCTATTTAAATTAATTTGGCCAGGTATAATG 2426
QY 2351 CCAGTCAGTCTCTTTATAGTGAAGAAATTTATTTGGTTAGTAAATATAATTTTAAACTA 2410
DB 2427 CCAGTCAGTCTCTTTATAGTGAAGAAATTTATTTGGTTAGTAAATATAATTTTAAACTA 2486
QY 2411 AATATATAATCTATTAATGTTTAAACATATGTTCAATTAAGCATAGCACTTTGAAATTA 2470
DB 2487 AATATATAATCTATTAATGTTTAAACATATGTTCAATTAAGCATAGCACTTTGAAATTA 2546
QY 2471 CTATATAATAGTCTATTTTACATTTACAGTCTTTTCAATTTGATGCTGCAATCTTT 2530
DB 2547 CTATATAATAGTCTATTTTACATTTACAGTCTTTTCAATTTGATGCTGCAATCTTT 2606
QY 2531 AGCATTAAAGGAAATGACTATGCTAATATTAATACCTGACCATGAAAAAATAAGTACCTC 2590
DB 2607 AGCATTAAAGGAAATGACTATGCTAATATTAATACCTGACCATGAAAAAATAAGTACCTC 2666
QY 2591 AAATGATGCAATTTGCACTGGTGAATCCAACTGCAAAATCTTTGTGCAATCTTTGATAT 2650
DB 2667 AAATGATGCAATTTGCACTGGTGAATCCAACTGCAAAATCTTTGTGCAATCTTTGATAT 2726
QY 2651 AGGTATTTTACATGGTGTGACATGACACACACCAATTTTCAATTTGATGACCTTTG 2710
DB 2727 AGGTATTTTACATGGTGTGACATGACACACCAATTTTCAATTTGATGACCTTTG 2786
QY 2711 AGGCTGTGCAATTTTCCACTTAAACCAACAGCTGAAAGTGAACCTCGAAACTTTGTT 2770
DB 2787 AGGCTGTGCAATTTTCCACTTAAACCAACAGCTGAAAGTGAACCTCGAAACTTTGTT 2846
QY 2771 TCATAAATCTTTCAAAAGTGTGTTTACATCAATGTTAAATTTCAAAATGCTGCAAGGTA 2830
DB 2847 TCATAAATCTTTCAAAAGTGTGTTTACATCAATGTTAAATTTCAAAATGCTGCAAGGTA 2906
QY 2831 ATTTAATGATATAATATTTAGTAAGAAAGTATGTTTGCATCTTATGTAAGTATGATC 2890
DB 2907 ATTTAATGATATAATATTTAGTAAGAAAGTATGTTTGCATCTTATGTAAGTATGATC 2966
QY 2891 ACAACATACAAATTTCAATTCAGTGCATGCTTTAGGTGTTAAGCATGAGATTTGATATGTT 2950
DB 2967 ACAACATACAAATTTCAATTCAGTGCATGCTTTAGGTGTTAAGCATGAGATTTGATATGTT 3026
QY 2951 TACTGTTAGTCTTGTGCACTCTGTTGCTAGTGAAGTATGAGAAAGTATGCAAGGACTGGA 3010
DB 3027 TACTGTTAGTCTTGTGCACTCTGTTGCTAGTGAAGTATGAGAAAGTATGCAAGGACTGGA 3086
QY 3011 CGTATTTTGTGCTTAAATAAAGGCTGTTTGTAGGCTTTTAAATATGCTTATTTTG 3070
DB 3087 CGTATTTTGTGCTTAAATAAAGGCTGTTTGTAGGCTTTTAAATATGCTTATTTTG 3146
QY 3071 TGTGCTCTCACTACTATTACACACTGTTGCTTTGTGCTTTGTTGTTGTTGTTGTTGTTGTTG 3130
DB 3147 TGTGCTCTCACTACTATTACACACTGTTGCTTTGTTGCTTTGTTGTTGTTGTTGTTGTTGTTG 3206
QY 3131 GTTATACAGTATGTTAAATTTCCATGAGAAAAATAATATGCTGTAATTTCTCAA 3184
DB 3207 GTTATACAGTATGTTAAATTTCCATGAGAAAAATAATATGCTGTAATTTCTCAA 3260
```

RESULT 8
ABV23741

ID ABV23741 standard; cDNA; 3498 BP.

XX

AC ABV23741;

XX

DT 16-SEP-2002 (first entry)

XX

```
DE Human prostate expression marker cDNA 23732.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
XX WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Mohahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 4356-4357; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification of its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 3498 BP; 1103 A; 609 C; 736 G; 1040 T; 0 U; 10 Other;
Query Match 98.3%; Score 3152.4; DB 5; Length 3498;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3167; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 11 GCGTCCGGCGGAGGACGAGCGGCGGCGGCGGCTCTGCGCGGCTCTCGCTCTC 70
DB 87 GAGCGCTGGCGAGGACGAGCGGCGGCGGCTCTGCGCGGCTCTCGCTCTC 146
QY 71 GGGCGGGCGGCGGCGGAGCGGAGCTCGGACTAGCGAACCCGAGGACGACATCAATAA 130
DB 147 GGGCGGGCGGCGGAGCGGAGCTCGGAGCTAGCGAACCCGAGGACGACATCAATAA 206
QY 131 ATAAATCATCAGAAATGACACCTTCTCAGGTTGCTTTGAAATTAAGAGGAATCTTTTAC 190
DB 207 ATAAATCATCAGAAATGACACCTTCTCAGGTTGCTTTGAAATTAAGAGGAATCTTTTAC 266
QY 191 CAGGAGAAATTTTGGCATATGTTGAGCTCTGATGCTTTGGGAACTGGGAATCTCTCAA 250
DB 267 CAGGAGAAATTTTGGCATATGTTGAGCTCTGATGCTTTGGGAACTGGGAATCTCTCAA 326
QY 251 ATGCTGTGGCTCTTCTTCCAGAGAAATGACACAGGTGAAAGCATCTATGGAAGCAACCA 310
DB 327 ATGCTGTGGCTCTTCTTCCAGAGAAATGACACAGGTGAAAGCATCTATGGAAGCAACCA 386
QY 311 TTGTACTCAGTAGAGGAGTATCAGTTTCAATTCAGTATCCCTACTTCAAAAGGCTACTTTTAGAAC 370
```

Db	387	TTGTACTCAGTAGAGGAGTATCAGTTTCTAGTATCGTACTCTTCAAAGGCTACTTTTTAGAAC	446
Qy	371	CAAAAGACTATCGGTGGTCCATGTCAAGTGTAGTGTCACAAGTGGGAGCTCATCTACAAC	430
Db	447	CAAAAGACTATCGGTGGTCCATGTCAAGTGTAGTGTCACAAGTGGGAGCTCATCTACAAC	506
Qy	431	CACGATCAATTAACCCCTTTAGAAAGCGAAATTTATTATTTAGCATGGACAATTTGGAATCC	490
Db	507	CACGATCAATTAACCCCTTTAGAAAGCGAAATTTATTATTTAGCATGGACAATTTGGAATCC	566
Qy	491	ACAAATGGTGTGAAACTCTGGATCTCGATGGCTGACATGTGACAGCTGAATAAGATTAC	550
Db	567	ACAAATGGTGTGAAACTCTGGATCTCGATGGCTGACATGTGACAGCTGAATAAGATTAC	626
Qy	551	GTTTGCATTATCTGAAAAAACCCTCTGTGTCAATAAACCAAGAAAAAATTTAAAAAATCTA	610
Db	627	GTTTGCATTATCTGAAAAAACCCTCTGTGTCAATAAACCAAGAAAAAATTTAAAAAATCTA	686
Qy	611	GATTTAGGGTGAAGCTGACACTAGNAGCCCTGGAGGAGATGACGATAGAGGTATCTC	670
Db	687	GATTTAGGGTGAAGCTGACACTAGNAGCCCTGGAGGAGATGACGATAGAGGTATCTC	746
Qy	671	CCACTGTACTCCAAAAATGTCCAAATAGCTTTGGAGATATCTTAAATAGCGACAATGAGT	730
Db	747	CCACTGTACTCCAAAAATGTCCAAATAGCTTTGGAGATATCTTAAATAGCGACAATGAGT	806
Qy	731	TCAAGTGCAGGCATTTCACAGCGGAGTGTGGTTATGGCTTCAGCCTGATCGTTGGACAG	790
Db	807	TCAAGTGCAGGCATTTCACAGCGGAGTGTGGTTATGGCTTCAGCCTGATCGTTGGACAG	866
Qy	791	AGTACAGCATACAGACGATGGAAACAGATTAACCTGGAACTAATCTTTGATTTTTGGAAG	850
Db	867	AGTACAGCATACAGACGATGGAAACAGATTAACCTGGAACTAATCTTTGATTTTTGGAAG	926
Qy	851	AGATCTCAGTGAGCACCTGAGTTTCAAGGTGATGCCCTTCCTGGACATGTGGGTACAGCTT	910
Db	927	AGATCTCAGTGAGCACCTGAGTTTCAAGGTGATGCCCTTCCTGGACATGTGGGTACAGCTT	986
Qy	911	GTCTCTTATCATCCACCAATGTGTAGAGTGGAAAGAGTGTGGAATCTTACTCTTCCCA	970
Db	987	GTCTCTTATCATCCACCAATGTGTAGAGTGGAAAGAGTGTGGAATCTTACTCTTCCCA	1046
Qy	971	TCATGAGCAGAAATTTCCCGGAAAAACAATAGGCAAAAGTGAGAGTTGACTATATAATTA	1030
Db	1047	TCATGAGCAGAAATTTCCCGGAAAAACAATAGGCAAAAGTGAGAGTTGACTATATAATTA	1106
Qy	1031	AGCCATTACCAGATACAGTTGTGACATGAATCTTCAATTTTCCAAAGTATTGGAGGCCAA	1090
Db	1107	AGCCATTACCAGATACAGTTGTGACATGAATCTTCAATTTTCCAAAGTATTGGAGGCCAA	1166
Qy	1091	GAATACCAATGGATGTGGCCATCGAGGTGCAGGAAACTCTACAAACAACCTGCCAGCTGG	1150
Db	1167	GAATACCAATGGATGTGGCCATCGAGGTGCAGGAAACTCTACAAACAACCTGCCAGCTGG	1226
Qy	1151	CTAAAGTTCAAGAAATACTATTTGCTCTTTAAGAAATGCTGTAGTCATGGTCAGCCT	1210
Db	1227	CTAAAGTTCAAGAAATACTATTTGCTCTTTAAGAAATGCTGTAGTCATGGTCAGCCT	1286
Qy	1211	TTGTAGAAATTTGACGTACACCTTTCAAGGACTTTGTGGCCGTGTATATCATGATCTTA	1270
Db	1287	TTGTAGAAATTTGACGTACACCTTTCAAGGACTTTGTGGCCGTGTATATCATGATCTTA	1346
Qy	1271	CCTGTGTTGTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTATTGAAATTC	1330
Db	1347	CCTGTGTTGTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTATTGAAATTC	1405
Qy	1331	CAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTTAAAGCTCACTCATGTGACTGCAC	1390
Db	1406	CAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTTAAAGCTCACTCATGTGACTGCAC	1465
Qy	1391	TGAAATCTTAAGGATCGAAAGAAATCTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAAAATC	1450
Db	1466	TGAAATCTTAAGGATCGAAAGAAATCTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAAAATC	1525

Qy	1451	AGCCATTTCTCTCTTAAGATGGTTTTAGAGTCTTTGCCAGAAAGATGTAGGTTTTAAACA	1510
Db	1526	AGCCATTTCTCTCTCTTAAGATGGTTTTAGAGTCTTTGCCAGAAAGATGTAGGTTTTAAACA	1585
Qy	1511	TTGAAATAAAATGGATCTCCAGCAAAAGGATGGAAATGGGATGGTAACTTATCAACAT	1570
Db	1586	TTGAAATAAAATGGATCTCCAGCAAAAGGATGGAAATGGGATGGTAACTTATCAACAT	1645
Qy	1571	ATTTTGACATGAATCTGTTTTTTGGATATAATTTTAAAAACTGTTTTAGAAAAATCTGGGA	1630
Db	1646	ATTTTGACATGAATCTGTTTTTTGGATATAATTTTAAAAACTGTTTTAGAAAAATCTGGGA	1705
Qy	1631	AGAGGAGATAGTGTGTTTTCTTCATTTTGATGCAGATATTTGCAAAATGGTTTCGGCAAAAGC	1690
Db	1706	AGAGGAGATAGTGTGTTTTCTTCATTTTGATGCAGATATTTGCAAAATGGTTTCGGCAAAAGC	1765
Qy	1691	AGAAACAAATATCCGATATATTTTTAACTCAAGGAAAAATCTGAGATTTATTCCTGAACTCA	1750
Db	1766	AGAAACAAATATCCGATATATTTTTAACTCAAGGAAAAATCTGAGATTTATTCCTGAACTCA	1825
Qy	1751	TGACCTCAGATCTCGGACAAACCCCAATTCGAATGAGCTTTGCACAGTTTGAANAATCTAC	1810
Db	1826	TGACCTCAGATCTCGGACAAACCCCAATTCGAATGAGCTTTGCACAGTTTGAANAATCTAC	1885
Qy	1811	TGGGGATAATGTACATCTGAAGACTTGTCTCAGAAACCCATCCTATATTTCAAGAGGCAAA	1870
Db	1886	TGGGGATAATGTACATCTGAAGACTTGTCTCAGAAACCCATCCTATATTTCAAGAGGCAAA	1945
Qy	1871	AAGCTAAGGACCTAGTCAATCTCTGCTGGGGTGAATACCAATGATCCTCGAAAAACAGAA	1930
Db	1946	AAGCTAAGGACCTAGTCAATCTCTGCTGGGGTGAATACCAATGATCCTCGAAAAACAGAA	2005
Qy	1931	GGAAATTTGAAGAACTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATGGATGC	1990
Db	2006	GGAAATTTGAAGAACTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATGGATGC	2065
Qy	1991	CTGAACCAACCAATATATTCCAAGTGGAGCAATTTGGAAACGCTTGAAGCAGAAATTCGCAG	2050
Db	2066	CTGAACCAACCAATATATTCCAAGTGGAGCAATTTGGAAACGCTTGAAGCAGAAATTCGCAG	2125
Qy	2051	AGCTTAAGACCTGTTGTGTCCACCTGTTAGCCGCTTTGTTCCCTCATCTTTGTGTGGGG	2110
Db	2126	AGCTTAAGACCTGTTGTGTCCACCTGTTAGCCGCTTTGTTCCCTCATCTTTGTGTGGGG	2185
Qy	2111	AGTCTGATATCCATGTGGATGCCAACCGCATGATAAACGTTGGAGAAATGCTTGTATTTAT	2170
Db	2186	AGTCTGATATCCATGTGGATGCCAACCGCATGATAAACGTTGGAGAAATGCTTGTATTTAT	2245
Qy	2171	TGACACAGAGTCAATTTTGGGGGGTGCACCGCTGTTCTGGGTATTCATTTTTCATCACTG	2230
Db	2246	TGACACAGAGTCAATTTTGGGGGGTGCACCGCTGTTCTGGGTATTCATTTTTCATCACTG	2305
Qy	2231	AGCAATGTTGATCTATGSCCTTTTGGGCTTCTCAGTTCAATGAAGCAATAATGAAGTATTT	2290
Db	2306	AGCAATGTTGATCTATGSCCTTTTGGGCTTCTCAGTTCAATGAAGCAATAATGAAGTATTT	2365
Qy	2291	AACTCTTTCTACTACAGTTCTTTGCAAGTATGCTATTTAAAAATTAATCTGGCCAGGATATTTG	2350
Db	2366	AACTCTTTCTACTACAGTTCTTTGCAAGTATGCTATTTAAAAATTAATCTGGCCAGGATATTTG	2425
Qy	2351	CCAGTCACTCTCTTTTATAGTGAGAAAAATTTATTTGGTTAGTAATAATAATTTTAAACTTA	2410
Db	2426	CCAGTCACTCTCTTTTATAGTGAGAAAAATTTATTTGGTTAGTAATAATAATTTTAAACTTA	2485
Qy	2411	AATATATAAATCTATATATGTTAAACATATCTTCAATTAABAGCATAGCACCTTTGAAATTA	2470
Db	2486	AATATATAAATCTATATATGTTAAACATATCTTCAATTAABAGCATAGCACCTTTGAAATTA	2545
Qy	2471	CTATATAAATAGCTCATATTTTACATTTACAGCTTTTTCATTTTGATCAGGCTCTGAAATCTTT	2530
Db	2546	CTATATAAATAGCTCATATTTTACATTTACAGCTTTTTCATTTTGATCAGGCTCTGAAATCTTT	2605

Db	487	TTCCTCTCTTTAAGAAATGCTGCTGATGCTGGTGCACCTTTTGTAGAAATTTGACGTACACC	546
Qy	1232	TTTTCAAAGGACTTTGTGCCGTGGTATATCATGATCTTACCTGTTGTTGACTATGAATAA	1291
Db	547	TTTTCAAAGGACTTTGTGCCGTGGTATATCATGATCTTACCTGTTGTTGACTATGAATAA	606
Qy	1292	AGAAATTTGATGCTGATCCAGTTTGAATTTATTTGAAATTTCCAGTAAAGAAATTAACATTTG	1351
Db	607	AGAAATTTGATGCTGATCCAGTTTGAATTTATTTGAAATTTCCAGTAAAGAAATTAACATTTG	666
Qy	1352	ACCAACTCCAGTTTGTAAAGCTCACTCATGTGACTGCACCTGAAATCTAAGGATCGAAAG	1411
Db	667	ACCAACTCCAGTTTGTAAAGCTCACTCATGTGACTGCACCTGAAATCTAAGGATCGAAAG	726
Qy	1412	AATCTGTGGTTTCAGGAGGAAATTTCTTTTCAGAAATCAGGCATTTCTTCTCTTAAGA	1471
Db	727	AATCTGTGGTTTCAGGAGGAAATTTCTTTTCAGAAATCAGGCATTTCTTCTCTTAAGA	786
Qy	1472	TGGTTTTAGAGTCTTTGCCAGAGAGATGTAGGGTTTTAACTTGAATTAATAATCGATCGC	1531
Db	787	TGGTTTTAGAGTCTTTGCCAGAGAGATGTAGGGTTTTAACTTGAATTAATAATCGATCGC	846
Qy	1532	AGCAAGGGGATGGAATGCTGGGATGGTTAACTTTATCAACATATTTTGACATGAATCTGTTTT	1591
Db	847	AGCAAGGGGATGGAATGCTGGGATGGTTAACTTTATCAACATATTTTGACATGAATCTGTTTT	906
Qy	1592	TGGATATAATTTTAAAAACTGTTTTTGAATAATTTCTGGGAAGAGGAAGATGTGTTTTCTT	1651
Db	907	TGGATATAATTTTAAAAACTGTTTTTGAATAATTTCTGGGAAGAGGAAGATGTGTTTTCTT	966
Qy	1652	CATTTGATGCAGATATTTTGACAAATGTTTCGGCAAAAGCAGAAATATCCGATACTAT	1711
Db	967	CATTTGATGCAGATATTTTGACAAATGTTTCGGCAAAAGCAGAAATATCCGATACTAT	1026
Qy	1712	TTTTTAACTCAAGGAAATCTGAGATTTTATCCTGAACTCATGAGACCTCAGATCTCGACAA	1771
Db	1027	TTTTTAACTCAAGGAAATCTGAGATTTTATCCTGAACTCATGAGACCTCAGATCTCGACAA	1086
Qy	1772	CCCCCATGCAATGAGCTTTTGCA CAGTTTGAAATCTACTGCGGATTAATGTTACATCTG	1831
Db	1087	CCCCCATGCAATGAGCTTTTGCA CAGTTTGAAATCTACTGCGGATTAATGTTACATCTG	1146
Qy	1832	AAAGCTTGCTCAGAAACCACTCTATATCAAGAGGCCAAAAGCTTAAGGACTAGTCAATAT	1891
Db	1147	AAAGCTTGCTCAGAAACCACTCTATATCAAGAGGCCAAAAGCTTAAGGACTAGTCAATAT	1206
Qy	1892	TCTGCTGGGGTGATGATACCAATGATCCTGAAAA CAGAAGGAAATTTGAAGGAACTTGGAG	1951
Db	1207	TCTGCTGGGGTGATGATACCAATGATCCTGAAAA CAGAAGGAAATTTGAAGGAACTTGGAG	1266
Qy	1952	TTAATGGTCTAAATTTATGATAGGATATATGATTTGGATGCCTGAAACCAAAATATATTC	2011
Db	1267	TTAATGGTCTAAATTTATGATAGGATATATGATTTGGATGCCTGAAACCAAAATATATTC	1326
Qy	2012	AAAGTGAGCAATTTGGAACGCTGAAGCAGGAAATTTGCCAGAGCTTAAGAGCTTTGTGTC	2071
Db	1327	AAAGTGAGCAATTTGGAACGCTGAAGCAGGAAATTTGCCAGAGCTTTGAAGGCTTTGTGTC	1386
Qy	2072	CCACTGTTTAGCGCTTTTCTCCCTCATCTTTGTGCGGAGTCTGATATCCATGTGGAATG	2131
Db	1387	CCACTGTTTAGCGCTTTTCTCCCTCATCTTTGTGCGGAGTCTGATATCCATGTGGAATG	1446
Qy	2132	CCAAAGCAATTTGATAAAGCTGGAAGATGCTTATGTTTTTATTCACAGAGGTCATTTTGGG	2191
Db	1447	CCAAAGCAATTTGATAAAGCTGGAAGATGCTTATGTTTTTATTCACAGAGGTCATTTTGGG	1506
Qy	2192	GGTGACCGCTGTTCTGGGTATTTCAATTTTTTCACTATGACATTTGTTGATATGCTT	2251
Db	1507	GGTGACCGCTGTTCTGGGTATTTCAATTTTTTCACTATGACATTTGTTGATATGCTT	1566
Qy	2252	TTGGGCTTCTCAGTTCAATGAAGCAATTAATGAAGTATTTTAACTCTTTTCACTACAGTTCTT	2311

Db	1567	TTGGGCTTCAGTTCAAATGAAGCAATAATGAAGTATTAACTCTTTCACTACAGTTCTT	1622
Qy	2312	GCAAGTATGCTATTTAAATTACTTGGCCAGGTATAAATGCCAGTCAGTCTCTTTATAGTG	2371
Db	1627	GCAAGTATGCTATTTAAATTACTTGGCCAGGTATAAATGCCAGTCAGTCTCTTTATAGTG	1686
Qy	2372	AGAAAATTTATTTGGTTAGPAATATAAATAATTTTAAACTAAATATATAAAATCTATAATGTT	2431
Db	1687	AGAAAATTTATTTGGTTAGTAATATAAATAATTTTAAACTAAATATATAAAATCTATAATGTT	1746
Qy	2432	AAACATATGTTCAATTTAAAGCATAGACATTTGAAATTAACATATATAAAATGACTCATATTT	2491
Db	1747	AAACATATGTTCAATTTAAAGCATAGACATTTGAAATTAACATATATAAAATGACTCATATTT	1806
Qy	2492	ACACTTTACAGCTTTTCATTTGATCAGCTCTGAAATCTTTAGCAGCTTAAAGGAAATGACATA	2551
Db	1807	ACACTTTACAGCTTTTCATTTGATCAGCTCTGAAATCTTTAGCAGCTTAAAGGAAATGACATA	1866
Qy	2552	TGCATAATTTATACCTGACCATGAAAAAAATAAGTACCTCAAATGCATGCAATTTGCACTGG	2611
Db	1867	TGCATAATTTATACCTGACCATGAAAAAAATAAGTACCTCAAATGCATGCAATTTGCACTGG	1926
Qy	2612	TGATTCGAAGTGCACAAATCTTTTGCCACTCTTGATATATAGGTATTTTTCACATGGTTCG	2671
Db	1927	TGATTCGAAGTGCACAAATCTTTTGCCACTCTTGATATATAGGTATTTTTCACATGGTTCG	1986
Qy	2672	ACATGCACACAAACACATTTTCATTCAGTATGAACCTTCAGGGCTGCTGCCATTTTTTCCAC	2731
Db	1987	ACATGCACACACACATTTTCATTCAGTATGAACCTTCAGGGCTGCTGCCATTTTTTCCAC	2046
Qy	2732	TTAACCAAAACGAGCTGGAAGGTGAACTCTGAAATCTTTTTCATAAATCTTTTCAAAAGTTG	2791
Db	2047	TTAACCAAAACGAGCTGGAAGGTGAACTCTGAAATCTTTTTCATAAATCTTTTCAAAAGTTG	2106
Qy	2792	TTTTACATCAATGTTTAAATTTCAAATGCTCGAGGTGAACTCTGAAATCTTTTCAAAAGTTG	2851
Db	2107	TTTTACATCAATGTTTAAATTTCAAATGCTCGAGGTGAACTCTGAAATCTTTTCAAAAGTTG	2166
Qy	2852	TAAGAAAAAGTATGTTTGCATCTTCTAGTAGAATAGATCAACATACAAATTCATTTCA	2911
Db	2167	TAAGAAAAAGTATGTTTGCATCTTCTAGTAGAATAGATCAACATACAAATTCATTTCA	2256
Qy	2912	GTGCATGCTTTAGTGTTTAAAGCATGAGATTTGTAAGTTTCTGTTTAGGTCCTTGCAATCT	2971
Db	2227	GTGCATGCTTTAGTGTTTAAAGCATGAGATTTGTAAGTTTCTGTTTAGGTCCTTGCAATCT	2286
Qy	2972	GTGCTGTAGTGATGATGAGAGATGTCGAAGCATGCTGGAAGTATTTGTTGCCCTAAAAAA	3031
Db	2287	GTGCTGTAGTGATGATGAGAGATGTCGAAGCATGCTGGAAGTATTTGTTGCCCTAAAAAA	2346
Qy	3032	AAAAGGCTGTTTGTAGGCGTTTTTAAATATGCTATTTTGTGTCTCTCACTACCTATTA	3091
Db	2347	AAAAGGCTGTTTGTAGGCGTTTTTAAATATGCTATTTTGTGTCTCTCACTACCTATTA	2406
Qy	3092	CACACTGTTGCTTTGTGGGTTTGTGTTGATGTCGGTGTGTTATACAGTAGTTAAATTC	3151
Db	2407	CACACTGTTGCTTTGTGGGTTTGTGTTGATGTCGGTGTGTTATACAGTAGTTAAATTC	2466
Qy	3152	CATGAGAAAAATAAATGCTCGAATTCCTCAA	3184
Db	2467	CATGAGAAAAATAAATGCTCGAATTCCTCAA	2499
RESULT 10			
ADJ57931 standard; cDNA; 2738 BP.			
XX	AC	ADJ57931;	
XX	XX	06-MAY-2004 (first entry)	
XX	DE	Human NARC 9 cDNA.	
XX	XX		

RESULT 10
ADJ57931
ID ADJ57931 standard; cDNA; 2738 BP.
XX
XX ADJ57931;
XX
XX
DT 06-MAY-2004 (first entry)
XX
XX Human NARC 9 cDNA.
XX

KW Neuronal apoptosis regulated candidate; NARC; diagnosis; cancer;
 KW lung disease; cirrhosis; hepatitis; atherosclerosis;
 KW myocardial infarction; inflammation; anaemia; glomerulonephritis;
 KW osteoporosis; AIDS; acquired immunodeficiency syndrome;
 KW Parkinson's disease; Alzheimer's disease; stroke; dermatitis;
 KW drug screening; gene therapy; cytostatic; hepatotropic; nootropic;
 KW cerebroprotective; dermatological; virucide; neuroprotective;
 KW phosphatidylglycerolphosphate synthase; PGP synthase; human; ss.
 XX Homo sapiens.
 XX US2004009553-A1.
 XX 15-JAN-2004.
 XX 30-APR-2003; 2003US-00426776.
 XX 27-SRP-1999; 99US-00406045.
 XX 22-OCT-1999; 99US-0161188P.
 XX 31-JAN-2000; 2000US-00495823.
 XX 28-FEB-2000; 2000US-0185517P.
 XX 20-OCT-2000; 2000US-00692785.
 XX 31-JAN-2001; 2001US-00773426.
 XX 28-FEB-2001; 2001US-00795691.
 XX 31-OCT-2001; 2001US-0335003P.
 XX 31-OCT-2001; 2001US-0335037P.
 XX 25-MAR-2002; 2002US-00105992.
 XX 28-AUG-2002; 2002US-00229662.
 XX 30-OCT-2002; 2002US-00284014.
 XX 30-OCT-2002; 2002US-00284059.
 XX 09-DEC-2002; 2002US-00314881.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Glucksmann MA, Williamson MJ, Tsai F, Rudolph-Owen LA;
 PI Kapeller-Libermann R, Meyers RE, Chiang LW, Hunter JJ;
 XX WPI; 2004-0904659/09.
 XX New isolated nucleic acid molecules and polypeptides (e.g. 27411, 23413,
 PT 22438, 23553, NARC SCI or NARC 1) useful for diagnosing, preventing or
 PT treating disorders associated with the protein, e.g. cancer,
 PT atherosclerosis or AIDS.
 XX Claim 1; SEQ ID NO 32; 260pp; English.
 XX The present invention provides isolated nucleic acid molecules and
 CC proteins designated 27411, 23413, 22438, 23553, 25278, NARC SCI,
 CC NARC 10A, NARC 1, NARC 12, NARC 13, NARC 17, NARC 25, NARC 3, NARC 4,
 CC NARC 7, NARC 8, NARC 11, NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20,
 CC NARC 26, NARC 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C,
 CC NARC 8B, NARC 9, NARC 2A, NARC 16B, NARC 1C, NARC 1A, NARC 25, 86604 or
 CC 32222. The invention is useful in diagnosing, preventing or treating
 CC disorders such as cancer, lung disease, cirrhosis, hepatitis,
 CC atherosclerosis, myocardial infarction, inflammation, anaemia,
 CC glomerulonephritis, osteoporosis, AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, stroke and
 CC dermatitis. These may also be used in drug screening. The invention is
 CC also useful in gene therapy. The present sequence is human neuronal
 CC apoptosis regulated candidate (NARC) cDNA.
 XX Sequence 2738 BP; 880 A; 450 C; 540 G; 868 T; 0 U; 0 Other;
 SQ
 Query Match 77.4%; Score 2481.8; DB 12; Length 2738;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2486; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 692 CCAATAGCTTGAGATATCTTAATAAGCGACAATGAGTCAAGTGCAGGCATTACACAGC 751
 DB 7 CCACGGCTCCGGAGATATCTTAATAAGCGACAATGAGTCAAGTGCAGGCATTACACAGC 66
 QY 752 CGAGTGTGGTATGGCTTGACGCTGATCGTTGACAGAGTACAGCATACAGAGTGG 811

Db 67 CGAGTGTGGTTATGGCTTTCGAGCCCTGATCGTGTGACAGAGTACAGCATACAGACGATGG 126
 QY 812 AACACAGATAAACCCTGGAACCTAATCTTTTGGATTTTTCGAAGAAGATCTCAGTGCAGCAGTAG 871
 Db 127 AACACAGATAAACCCTGGAACCTAATCTTTTGGATTTTTCGAAGAAGATCTCAGTGCAGCAGTAG 186
 QY 872 TTCAGGGTGATGCCCTTCTCTGGAACATGTTGGGTACAGCTTGTCTCTTATCATCCACATTTG 931
 Db 187 TTCAGGGTGATGCCCTTCTCTGGAACATGTTGGGTACAGCTTGTCTCTTATCATCCACATTTG 246
 QY 932 CTGAGAGTGGAAAGAGTGTCTGGAATTTCTTCTTCCCATCATGAGCAGAAATTTCCCGGA 991
 Db 247 CTGAGAGTGGAAAGAGTGTCTGGAATTTCTTCTTCCCATCATGAGCAGAAATTTCCCGGA 306
 QY 992 AAACAATAGGCAAGTGTGAGAGTTGACTATATAATTAATTAAGCCATTTACAGGATACAGTT 1051
 Db 307 AAACAATAGGCAAGTGTGAGAGTTGACTATATAATTAATTAAGCCATTTACAGGATACAGTT 366
 QY 1052 GTGACATGAATCTTCAATTTTCCAAAGTATTTGGAAGCCAAAGATACCATTTGGATTTGGCC 1111
 Db 367 GTGACATGAATCTTCAATTTTCCAAAGTATTTGGAAGCCAAAGATACCATTTGGATTTGGCC 426
 QY 1112 ATCGAGGTGCAGGAACCTCTACCAACTGCCAGCTGGCTAAAGTTCAAGAAATACATA 1171
 Db 427 ATCGAGGTGCAGGAACCTCTACCAACTGCCAGCTGGCTAAAGTTCAAGAAATACATA 486
 QY 1172 TTGCTTTCTTTAAGAAATGCTGTAGTTCATGCTGAGCCCTTTGTAGAAATTTGACGTACAC 1231
 Db 487 TTGCTTTCTTTAAGAAATGCTGTAGTTCATGCTGAGCCCTTTGTAGAAATTTGACGTACAC 546
 QY 1232 TTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTGTTGTGACTTGAAGA 1291
 Db 547 TTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTGTTGTGACTTGAAGA 606
 QY 1292 AGAAATTTGATGCTGATCCAGTTCGAATTTTGAATTTCCAGTAAAGAAATTAACATTTG 1351
 Db 607 AGAAATTTGATGCTGATCCAGTTCGAATTTTGAATTTCCAGTAAAGAAATTAACATTTG 666
 QY 1352 ACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTGAAATCTAAAGATCGGAAG 1411
 Db 667 ACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTGAAATCTAAAGATCGGAAG 726
 QY 1412 AATCTGTGGTTCAGAGGAAATTTCTTTTTCAGAAATTCAGCCATTTCTCTCTTAAGA 1471
 Db 727 AATCTGTGGTTCAGAGGAAATTTCTTTTTCAGAAATTCAGCCATTTCTCTCTTAAGA 786
 QY 1472 TGGTTTATAGCTCTTTGCCAGAAATGTAGGGTTTAAATTTGAATTTAAATTTGATCTGCC 1531
 Db 787 TGGTTTATAGCTCTTTGCCAGAAATGTAGGGTTTAAATTTGAATTTAAATTTGATCTGCC 846
 QY 1532 AGCAAGGGATGGAATGTGGATGTTAACTTATCAACATATTTTGAATTTGATCTGTTT 1591
 Db 847 AGCAAGGGATGGAATGTGGATGTTAACTTATCAACATATTTTGAATTTGATCTGTTT 906
 QY 1592 TGGATATAATTTTAAAACTGTTTGAATAATTTGGAAGAGGAGAAATAGTGTGTTTCTT 1651
 Db 907 TGGATATAATTTTAAAACTGTTTGAATAATTTGGAAGAGGAGAAATAGTGTGTTTCTT 966
 QY 1652 CATTTGATGAGATATTTTGCACATGTTTGGCAAAAGCAGAAATATCCGATCTAT 1711
 Db 967 CATTTGATGAGATATTTTGCACATGTTTGGCAAAAGCAGAAATATCCGATCTAT 1026
 QY 1712 TTTTAACTCAAGGAAATCTGAGATTTTCTGAACTTCATGACCTCAGATCTCGGACAA 1771
 Db 1027 TTTTAACTCAAGGAAATCTGAGATTTTCTGAACTTCATGACCTCAGATCTCGGACAA 1086
 QY 1772 CCCCATTGCAATGAGCTTTTGCACAGTTTGAATAATTTCTGCGGAGTAAATGTACATCTG 1831
 Db 1087 CCCCATTGCAATGAGCTTTTGCACAGTTTGAATAATTTCTGCGGAGTAAATGTACATCTG 1146
 QY 1832 AAGACTTGCTCAGAAACCCATCTTATTTCAAGAGGCAAAAGCTAAGGACTAGTCTAT 1891
 Db 1147 AAGACTTGCTCAGAAACCCATCTTATTTCAAGAGGCAAAAGCTAAGGACTAGTCTAT 1206

CC suitable for binding of the test compound to the polypeptide. The method
CC of the invention has haematological and cytostatic applications and may
CC be useful for identifying compounds for treating a haematological
CC disorder associated with erythroid cells e.g. anaemia and erythrocytosis,
CC bone marrow e.g. leukaemia, platelets e.g. thrombocytopenia and
CC thrombosis or B-cells and T-cells e.g. neutropenia. The compounds
CC identified may be utilised during gene therapy procedures. The current
CC sequence is that of a human haematological disorder-related cDNA of the
CC invention.

XX
SQ Sequence 2019 BP; 633 A; 363 C; 449 G; 574 T; 0 U; 0 Other;

Query Match 63.0%; Score 2019; DB 13; Length 2019;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	145	ATGACACCTTCTCAGGTGGCTTTGAAATAAGAGGAACCTCTTTACCAAGGAGAGTTT	204
DB	1	ATGACACCTTCTCAGGTGGCTTTGAAATAAGAGGAACCTCTTTACCAAGGAGAGTTT	60
QY	205	GGGATATGGAGCTGTGATCTTTGGGAACTCGGAATCTCAAAATCTGTGGCTCTT	264
DB	61	GGGATATGGAGCTGTGATCTTTGGGAACTCGGAATCTCAAAATCTGTGGCTCTT	120
QY	265	CTTCCAGAGATGACACAGGTGAAGCATGCTATGGAAGCAACCAATGTAATCTAGTAGA	324
DB	121	CTTCCAGAGATGACACAGGTGAAGCATGCTATGGAAGCAACCAATGTAATCTAGTAGA	180
QY	325	GGAGTATCAGTTCAGTATCGTACTTTCAAGGGTACTTTTGAACCAAGACTATCGGT	384
DB	181	GGAGTATCAGTTCAGTATCGTACTTTCAAGGGTACTTTTGAACCAAGACTATCGGT	240
QY	385	GGTCCATGTCAAGTGTAGTTCACAGTGGGAGACTCATCTCAACACGATCAATAACC	444
DB	241	GGTCCATGTCAAGTGTAGTTCACAGTGGGAGACTCATCTCAACACGATCAATAACC	300
QY	445	CCTTTAGAAAGCGAAATTAATATGACGATGACAAATTTGGAATCCCAATGGTGTGAA	504
DB	301	CCTTTAGAAAGCGAAATTAATATGACGATGACAAATTTGGAATCCCAATGGTGTGAA	360
QY	505	ACTCTGGATCTCGGAGGCTGACATGTCAGACTGCAATTAAGATTAAGTTTCATTTCT	564
DB	361	ACTCTGGATCTCGGAGGCTGACATGTCAGACTGCAATTAAGATTAAGTTTCATTTCT	420
QY	565	GAAAAACCTCTCTGTCAATAACCAAGAAAAATTTAAAAAATCTAGATTTAGGGTGAAG	624
DB	421	GAAAAACCTCTCTGTCAATAACCAAGAAAAATTTAAAAAATCTAGATTTAGGGTGAAG	480
QY	625	CTGACATAGAGGCTCGGAGGAAGATGACGATAGAGGTATCTCCCACTGTACTCCAC	684
DB	481	CTGACATAGAGGCTCGGAGGAAGATGACGATAGAGGTATCTCCCACTGTACTCCAC	540
QY	685	AAATGTCCATAGCTTGAGATATCTTAATAGCGCAATGAGTTCAAGTTCAGGCTAT	744
DB	541	AAATGTCCATAGCTTGAGATATCTTAATAGCGCAATGAGTTCAAGTTCAGGCTAT	600
QY	745	TCACAGCCGGAGTGTGGTTATGGCTTGAGCCCTGATCGTTGGACAGATPACAGCATACAG	804
DB	601	TCACAGCCGGAGTGTGGTTATGGCTTGAGCCCTGATCGTTGGACAGATPACAGCATACAG	660
QY	805	ACGATGGAAACCAAGTAACTCGGAATCTTTGATTTTTTTCGAAAGAGATCTCAGTAGAG	864
DB	661	ACGATGGAAACCAAGTAACTCGGAATCTTTGATTTTTTTCGAAAGAGATCTCAGTAGAG	720
QY	865	CAGGTAGTTTCAGGCTGATGCCCTTCTCGACATGTTGGGTACAGCTGTCTTATCATCC	924
DB	721	CAGGTAGTTTCAGGCTGATGCCCTTCTCGACATGTTGGGTACAGCTGTCTTATCATCC	780
QY	925	ACCATTTCTGAGATGGAAGAGTGTGGAAATCTTACTCTTTCCCATCATGAGCAGAAAT	984
DB	781	ACCATTTCTGAGATGGAAGAGTGTGGAAATCTTACTCTTTCCCATCATGAGCAGAAAT	840
QY	985	TCCCGGAAAAAATAAGGCAAGGTGAGTTGACTATATAATTAATTAAGCCATTACCAGGA	1044

DB	841	TCCCGGAAAAAATAAGGCAAGGTGAGATTGACTATATAATTAATTAAGCCATTACCAGGA	900
QY	1045	TACAGTTGTGACATCAAAATCTTCAATTTTCAAGTATTCGAAGCCCAAGAAATACCATTTGAT	1104
DB	901	TACAGTTGTGACATCAAAATCTTCAATTTTCAAGTATTCGAAGCCCAAGAAATACCATTTGAT	960
QY	1105	GTGCGCCATCGAGGTGCGAGGAAACTCTACAACTGCCAGCTGGCTTAAAGTTCAAGAA	1164
DB	961	GTGCGCCATCGAGGTGCGAGGAAACTCTACAACTGCCAGCTGGCTTAAAGTTCAAGAA	1020
QY	1165	ATATCTATTGCTTCTTTAAGAAATGCTGCTAGTCAATGCTGAGCCCTTTGTGAATTTGAC	1224
DB	1021	ATATCTATTGCTTCTTTAAGAAATGCTGCTAGTCAATGCTGAGCCCTTTGTGAATTTGAC	1080
QY	1225	GTACACCTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTTGTTGACT	1284
DB	1081	GTACACCTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTTGTTGACT	1140
QY	1285	ATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTATTTGAAATTCAGTAAAGAAATTA	1344
DB	1141	ATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTATTTGAAATTCAGTAAAGAAATTA	1200
QY	1345	ACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGCTGCACTGCAATCTAAGAT	1404
DB	1201	ACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGCTGCACTGCAATCTAAGAT	1260
QY	1405	CGGAAAGAAATCTGTGGTTTCAAGAGGAAATTTCTTTTCAGAAAAATCAGCCATTTCTTCT	1464
DB	1261	CGGAAAGAAATCTGTGGTTTCAAGAGGAAATTTCTTTTCAGAAAAATCAGCCATTTCTTCT	1320
QY	1465	CTTAAGATGGTTTTAGAGTCTTTGCGCAGAAGATGTAGGGTTTTAATTAATTAATTAATTAAT	1524
DB	1321	CTTAAGATGGTTTTAGAGTCTTTGCGCAGAAGATGTAGGGTTTTAATTAATTAATTAATTAAT	1380
QY	1525	ATCTGCCAGAAAGGGATGGAAATGTGGATGGTAACTTTATCAACATATTTTGTGACATGAAT	1584
DB	1381	ATCTGCCAGAAAGGGATGGAAATGTGGATGGTAACTTTATCAACATATTTTGTGACATGAAT	1440
QY	1585	CTGTTTTTGGATATAATTTTAAAGAACTGTTTTAGAAAAATCTGGGAAGAGGAGATAGTG	1644
DB	1441	CTGTTTTTGGATATAATTTTAAAGAACTGTTTTAGAAAAATCTGGGAAGAGGAGATAGTG	1500
QY	1645	TTTTCTTCAATTTGATGACAGATATTTGCACAATGGTTTGGCAAAAAGCAGAAATATCCG	1704
DB	1501	TTTTCTTCAATTTGATGACAGATATTTGCACAATGGTTTGGCAAAAAGCAGAAATATCCG	1560
QY	1705	ATACTATTTTAACTCAAGGAAATCTGAGATTTTATCTGAACTCATGGAATCTCAGATCT	1764
DB	1561	ATACTATTTTAACTCAAGGAAATCTGAGATTTTATCTGAACTCATGGAATCTCAGATCT	1620
QY	1765	CGGCAACCCCATTTGCAATGAGCTTTGCAAGTTTGAATAATCTACTGGGGATATAATGTA	1824
DB	1621	CGGCAACCCCATTTGCAATGAGCTTTGCAAGTTTGAATAATCTACTGGGGATATAATGTA	1680
QY	1825	CATATCTGAAGACTTGTCTCAGAAACCCATCTTATATTTCAAGAGGCAAAAGCTTAAGGACTA	1884
DB	1681	CATATCTGAAGACTTGTCTCAGAAACCCATCTTATATTTCAAGAGGCAAAAGCTTAAGGACTA	1740
QY	1885	GTCAATTTCTGCTGGGTGATGATACCAATGATCTGTAATAACAGAGGAAATTTGAAGAA	1944
DB	1741	GTCAATTTCTGCTGGGTGATGATACCAATGATCTGTAATAACAGAGGAAATTTGAAGAA	1800
QY	1945	CTTGGAGTTAATGCTTAATTTATGATAGGATATATGATTTGATGGATGCTTGAACCAAT	2004
DB	1801	CTTGGAGTTAATGCTTAATTTATGATAGGATATATGATTTGATGGATGCTTGAACCAAT	1860
QY	2005	ATATTTCAAAGTGGAGCAATTTGAAACCGCTGAAGCAGGAAATTTGCCAGAGCTTTAAGAGCTGT	2064
DB	1861	ATATTTCAAAGTGGAGCAATTTGAAACCGCTGAAGCAGGAAATTTGCCAGAGCTTTAAGAGCTGT	1920
QY	2065	TTGTGCTCCACTGTGTAGCCGCTTTGTTCCTCATCTTTTGTGTGGGAGTCTGTATCCAT	2124

Db 1921 TTGTGTCCTCACTGTAGCGCTTTGTTCCCTCATCTTTGTGTGGGAGTCTGATATCCAT 1980

QY 2125 GTGGATGCCAACGGCAATGTATAAAGTGGAGAAATGCTTAG 2163

Db 1981 GTGGATGCCAACGGCAATGTATAAAGTGGAGAAATGCTTAG 2019

RESULT 12

AAH14905

ID AAH14905 standard; cDNA; 1803 BP.

AC AAH14905;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human cDNA sequence SEQ ID NO:12778.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-00116126.

XX

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX

PA (HELI-) HELIX RES INST.

XX

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX

PS Claim 8; SEQ ID NO 12778; 2537pp + Sequence Listing; English.

XX

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and, in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX

SQ Sequence 1803 BP; 561 A; 291 C; 355 G; 596 T; 0 U; 0 Other;

Query Match 56.2%; Score 1803; DB 4; Length 1803;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 1803; Conservative 0;

QY 1379 ATGTGACTGCACCTGAAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTCCT 1438

Db 1 ATGTGACTGCACCTGAAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTCCT 60

QY 1439 TTTCAGAAATCAGCCATTTCTTCTCTTAAGATGTTTTAGAGTCTTTTGCCAGAAAGATG 1498

Db 61 TTTCAGAAATCAGCCATTTCTTCTCTTAAGATGTTTTAGAGTCTTTTGCCAGAAAGATG 120

QY 1499 TAGGGTTTAAACATTGAAATAAATAATGGATCTGCCAGCAAAAGGGATGGAATGTGGATGGTA 1558

Db 121 TAGGGTTTAAACATTGAAATAAATAATGGATCTGCCAGCAAAAGGGATGGAATGTGGATGGTA 180

QY 1559 ACTTATCAACATATTTTGACATGAATCTGTTTTTTGGATATAAATTTTAAAACTGTTTTAG 1618

Db 181 ACTTATCAACATATTTTGACATGAATCTGTTTTTTGGATATAAATTTTAAAACTGTTTTAG 240

QY 1619 AAAATTTCTGGAGAGGAGGAGATAGTTTTCTTCATTTGATCGAGATATTTGACAAATGG 1678

Db 241 AAAATTTCTGGAGAGGAGGAGATAGTTTTCTTCATTTGATCGAGATATTTGACAAATGG 300

QY 1679 TTGGCAAAAGCAGAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTT 1738

Db 301 TTGGCAAAAGCAGAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTT 360

QY 1739 ATCTGAACTCATGGACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTGCAAGT 1798

Db 361 ATCTGAACTCATGGACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTGCAAGT 420

QY 1799 TTGAAATCTACTGGGGATAAATGTACATCTGAAAGACTTGTCTCAGAAACCCCATCTATA 1858

Db 421 TTGAAATCTACTGGGGATAAATGTACATCTGAAAGACTTGTCTCAGAAACCCCATCTATA 480

QY 1859 TTCAAGAGGCAAAAGCTTAAGGACTAGTCTATTTCTGCTGGGTGATGATACCAATGATC 1918

Db 481 TTCAAGAGGCAAAAGCTTAAGGACTAGTCTATTTCTGCTGGGTGATGATACCAATGATC 540

QY 1919 CTGAAACAGAGGAAATTTGAAGGAACTTGGAGTTAATGGTCTAATTTATGATAGGATAT 1978

Db 541 CTGAAACAGAGGAAATTTGAAGGAACTTGGAGTTAATGGTCTAATTTATGATAGGATAT 600

QY 1979 ATGATGGATGCTGAAACCAAAATATATTCAGTGGAGGAAATTTGGAACGCTGAGC 2038

Db 601 ATGATGGATGCTGAAACCAAAATATATTCAGTGGAGGAAATTTGGAACGCTGAGC 660

QY 2039 AGGAATTCAGAGCTTAAGAGCTGTTGTGCTCCACTGTTAGCCGCTTTTGTTCCTCAT 2098

Db 661 AGGAATTCAGAGCTTAAGAGCTGTTGTGCTCCACTGTTAGCCGCTTTTGTTCCTCAT 720

QY 2099 CTTTGTGTGGGAGTCTGATATCCATGTGATGCCAACGGCAATTTGATGAGGAAATG 2158

Db 721 CTTTGTGTGGGAGTCTGATATCCATGTGATGCCAACGGCAATTTGATGAGGAAATG 780

QY 2159 CTTAGTTTTTATTTGACAGAGGTCATTTTGGGGCGTGCACCGCTGTTCTGGGTATTCAT 2218

Db 781 CTTAGTTTTTATTTGACAGAGGTCATTTTGGGGCGTGCACCGCTGTTCTGGGTATTCAT 840

QY 2219 TTTTCATCAGTGGATTTGATCTATGCTTTTGGGCTTCTCAGTTCAATGAGCAAT 2278

Db 841 TTTTCATCAGTGGATTTGATCTATGCTTTTGGGCTTCTCAGTTCAATGAGCAAT 900

QY 2279 AATCAAGTATTTAACTCTTTTCACTACAGTTTCTTGCAGATGATGCTATTTAAATTTGGC 2338

Db 901 AATCAAGTATTTAACTCTTTTCACTACAGTTTCTTGCAGATGATGCTATTTAAATTTGGC 960

QY 2339 CAGGTATAATTTGCCAGTCAGTCTCTTTTATAGTGAGAAATTTATTTGTTAGTAAATATAA 2398

Db 961 CAGGTATAATTTGCCAGTCAGTCTCTTTTATAGTGAGAAATTTATTTGTTAGTAAATATAA 1020

QY 2399 TATTTTAACTAAATATATAAATCTAATGTTTAAACATATGTTTCAATTAAGCATAGCA 2458

QY 1950 AGTTAATGGTCTAATTTATGATAGGATATATGATTGGATGCTTGAACAAACCAATATATT 2009
Db |
481 AGTTAATGGTCTAATTTATGATAGGATATATGATTGGATGCTTGAACAAACCAATATATT 540
QY 2010 CCAAGTGGAGCAATTGGAAACCCCTGAAGCAGCAATTCGACAGCTTAAGAGCTGTTTGTG 2069
Db |
541 CCAAGTGGAGCAATTGGAAACCCCTGAAGCAGCAATTCGACAGCTTAAGAGCTGTTTGTG 600
QY 2070 TCCCACTGTTAGCCGCTTTGTTCCCTCATCTTTGTGGGAGTCTGATATCCATGTGGA 2129
Db |
601 TCCCACTGTTAGCCGCTTTGTTCCCTCATCTTTGTGGGAGTCTGATATCCATGTGGA 660
QY 2130 TCCCAACGGCATTTGATACGTTGGAGAACTGCTAGTTTATTTGACAGAGTCAATTTGG 2189
Db |
661 TCCCAACGGCATTTGATACGTTGGAGAACTGCTAGTTTATTTGACAGAGTCAATTTGG 720
QY 2190 GGGCGTGACCGCTGCTGCTGGGTATTTCAATTTTCACTACGCAATTTGTCATCTATGCC 2249
Db |
721 GGGCGTGACCGCTGCTGCTGGGTATTTCAATTTTCACTACGCAATTTGTCATCTATGCC 780
QY 2250 TTTTGGGCTTCTCAGTTCAATGAAGCAATAATGAAGTATTTAACTCTTTCACTACAGTTC 2309
Db |
781 TTTTGGGCTTCTCAGTTCAATGAAGCAATAATGAAGTATTTAACTCTTTCACTACAGTTC 840
QY 2310 TTGCAAGTATGCTATTTAAATTTACTTGGCCAGTATTAATGCCAGTCACTCTTTATAG 2369
Db |
841 TTGCAAGTATGCTATTTAAATTTACTTGGCCAGTATTAATGCCAGTCACTCTTTATAG 900
QY 2370 TCAGAAATTTATGCTAGTAAATAATAATTTTAACTAAATATAATAATAATAATAATAATG 2429
Db |
901 TCAGAAATTTATGCTAGTAAATAATAATTTTAACTAAATATAATAATAATAATAATAATG 960
QY 2430 TTAACATATGTTCAATTAAGCATAGCACTTTGAAATTTAACTATATAAATAGTCTCATAT 2489
Db |
961 TTAACATATGTTCAATTAAGCATAGCACTTTGAAATTTAACTATATAAATAGTCTCATAT 1020
QY 2490 TTACACTTACAGCTTTTCAATTTGATCAGCTCGAAATCTTTAGCACTTAAGGAAATGAC 2549
Db |
1021 TTACACTTACAGCTTTTCAATTTGATCAGCTCGAAATCTTTAGCACTTAAGGAAATGAC 1080
QY 2550 TATGCATATTTATCTGACCATGAAATAAATAGTCTCAAAATGATGATTTGCACT 2609
Db |
1081 TATGCATATTTATCTGACCATGAAATAAATAGTCTCAAAATGATGATTTGCACT 1140
QY 2610 GGTGATTTCAACTGCAAAATCTTTGTGCCATCTTTGTATATAGGTAATTTTACATGGT 2669
Db |
1141 GGTGATTTCAACTGCAAAATCTTTGTGCCATCTTTGTATATAGGTAATTTTACATGGT 1200
QY 2670 TGAATGCAACACCACTTTTCAATTCAGTATGAACCTTGAGGCTGCTGCCATTTTCC 2729
Db |
1201 TGGCATGCAACACCACTTTTCAATTCAGTATGAACCTTGAGGCTGCTGCCATTTTCC 1260
QY 2730 ACTTAAACCAACAGGCTGAAGTGAACCTTGAACCTTTTCAATAAATCTTTCAAAAGT 2789
Db |
1261 ACTTAAACCAACAGGCTGAAGTGAACCTTGAACCTTTTCAATAAATCTTTCAAAAGT 1320
QY 2790 TGTGTTTACATCAATGTTTAAATTTCAAAATGCTGAGGCTAAATTTAATGATAAATAATTT 2849
Db |
1321 TGTGTTTACATCAATGTTTAAATTTCAAAATGCTGAGGCTAAATTTAATGATAAATAATTT 1380
QY 2850 AGTAAGAAAAAGTATGTTATGCACTACTAGTAAATAGATCAACATACAAATCAATTT 2909
Db |
1381 AGTAAGAAAAAGTATGTTATGCACTACTAGTAAATAGATCAACATACAAATCAATTT 1440
QY 2910 CAGTGAATGCTTTAGTGTTAAGCATGAGATTGATGCTTACTGTTAGTCTGCTGCTGAT 2969
Db |
1441 CAGTGAATGCTTTAGTGTTAAGCATGAGATTGATGCTTACTGTTAGTCTGCTGCTGAT 1500
QY 2970 CTGCTGCTAGTGTAGTATGAGAAGATGTCAGGACTGAGCGTATTTTGTGCTCAAAAT 3029
Db |
1501 CTGCTGCTAGTGTAGTATGAGAAGATGTCAGGACTGAGCGTATTTTGTGCTCAAAAT 1559
QY 3030 AAAAAAGGCTGTTTGTAGGCGTTTAAATAATGCTTATTTTGTGCTCTCTCACTACCTAT 3089

Db |
1560 AAAAAAGGCTGTTTGTAGGCGTTTAAATATGCTTATTTTGTGCTCTCTCACTACCTAT 1619
QY 3090 TACACACTGTTGCTTTGTTGGGTTTGTGTTATGTCGCTGTTATACAGTAGTTAAATTT 3149
Db |
1620 TACACACTGTTGCTTTGTTGGGTTTGTGTTATGTCGCTGTTATACAGTAGTTAAATTT 1679
QY 3150 TCCATGTCAGAAAAATAAATGCTCTGAAATTTCTCAA 3184
Db |
1680 TCCATGTCAGAAAAATAAATGCTCTGAAATTTCTCAA 1714
RESULT 14
ADJ57938
ID ADJ57938 standard; cDNA; 3381 BP.
XX
AC ADJ57938;
XX
XX
DT 06-MAY-2004 (first entry)
XX
XX Rat NARC 16 cDNA.
XX
XX Neuronal apoptosis regulated candidate; NARC; diagnosis; cancer;
KW lung disease; cirrhosis; hepatitis; atherosclerosis;
KW myocardial infarction; inflammation; anaemia; glomerulonephritis;
KW osteoporosis; AIDS; acquired immunodeficiency syndrome;
KW Parkinson's disease; Alzheimer's disease; stroke; dermatitis;
KW drug screening; gene therapy; cytostatic; hepatotropic; nootropic;
KW cerebroprotective; dermatological; virucide; neuroprotective;
KW phosphatidyglycerolphosphate synthase; PGP synthase; rat; ss.
XX
XX Rattus norvegicus.
XX
XX US2004009553-A1.
XX
XX 15-JAN-2004.
XX
XX 30-APR-2003; 2003US-00426776.
XX
XX 27-SEP-1999; 99US-00406045.
XX 22-OCT-1999; 99US-0161188P.
XX 31-JAN-2000; 2000US-00495823.
XX 28-FEB-2000; 2000US-0185517P.
XX 20-OCT-2000; 2000US-00692785.
XX 31-JAN-2001; 2001US-00773426.
XX 28-FEB-2001; 2001US-00795691.
XX 31-OCT-2001; 2001US-0335003P.
XX 25-MAR-2002; 2002US-00105992.
XX 28-AUG-2002; 2002US-00229662.
XX 30-OCT-2002; 2002US-00284014.
XX 30-OCT-2002; 2002US-00284059.
XX 09-DEC-2002; 2002US-00314881.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Gluckmann MA, Williamson MJ, Tsai F, Rudolph-Owen LA;
PI Kapeller-Libermann R, Meyers RE, Chiang LW, Hunter JU;
XX
XX WPI; 2004-090469/09.
XX
XX New isolated nucleic acid molecules and polypeptides (e.g. 27411, 23413,
PT 22438, 23553, NARC SCI or NARC 1) useful for diagnosing, preventing or
PT treating disorders associated with the protein, e.g. cancer,
PT atherosclerosis or AIDS.
XX
XX Claim 1; SEQ ID NO 39; 260pp; English.
XX
XX The present invention provides isolated nucleic acid molecules and
CC proteins designated 27411, 23413, 22438, 23553, 25278, 26212, NARC SCI,
CC NARC 10A, NARC 1, NARC 12, NARC 13, NARC 17, NARC 25, NARC 3, NARC 4,
CC NARC 7, NARC 8, NARC 13, NARC 14, NARC 15, NARC 16, NARC 19, NARC 20,
CC NARC 26, NARC 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C,

CC	NARC 8B, NARC 9, NARC 2A, NARC 16B, NARC 1C, NARC 1A, NARC 25, 86604 or 32222. The invention is useful in diagnosing, preventing or treating disorders such as cancer, lung diseases, cirrhosis, hepatitis, atherosclerosis, myocardial infarction, inflammation, anaemia, glomerulonephritis, osteoporosis, AIDS (acquired immunodeficiency syndrome), Parkinson's disease, Alzheimer's disease, stroke and dermatitis. These may also be used in drug screening. The invention is also useful in gene therapy. The present sequence is rat neuronal apoptosis regulated candidate (NARC) cDNA.	XX
CC	Sequence 3381 BP; 990 A; 689 C; 729 G; 973 T; 0 U; 0 Other;	
Query Match	45.9%; Score 1471.2; DB 12; Length 3381;	
Best Local Similarity	86.7%; Pred. No. 0;	
Matches 1671; Conservative	0; Mismatches 243; Indels 14; Gaps 4	
QY	55 CGGGTCTCGCTCTCGGCGGGCGCGCGAGC--GGACCTGCGGACTAGCGAACC 112	
DB	94 CGAGGCGGACATCTCGGCTTAGTCGCGCGCGGCCAGGCTCCGCGCGCGCGGACG 153	
QY	113 GGAGCA CGA CAT CATAAAATAAATCCATCAGAATGACACCTTCTCAGGTGGCCTTTGAAA 172	
DB	154 GAGTGGCAGAAATCTTAAATAATCCATCAGAATGACACCTTCTCAGGTCAACCTTTGAAA 213	
QY	173 TAAGAGGAATCTTTTACAGAGAAAGTTTTTGGATATGTGGAAGCTGTGATGCTTTGG 232	
DB	214 TAAGAGGAATCTTTTACAGAGAGAGTCTTTGCAATGTGTGAAACATGTGATGCCCTTGG 273	
QY	233 GAAACTGGAAATCCTCAAAATGCTGGCTCTTCTTCCAGAGAGATGACACAGGTGGAAGCA 292	
DB	274 GAAACTGGAGTCTCTCAAAATGCTGTG---CCTCTTACTGAGAGTGAGACAGGCGAAAG-- 328	
QY	293 TGTATGGAAGCAACCATTTGACTCAGTAGAGGAGTATCAGTTTCAGTATCGCTACTTCA 352	
DB	329 -TGTATGGAAGCAGTGATTTGTTCTTAGTAGGGAATGTCGTGAAGTACCGCTACTTCA 387	
QY	353 AAGGTGATCTTTTAGAACCAAAAGACTATCGGTGTCATGTCAGTGATAGTTTCAAGT 412	
DB	388 GAGGCTGCTTTTTAGAACCAAAAGACTATCGGTGGTCCATGTCAAGTCATAGTTTCAAGT 447	
QY	413 GGGAGACTCATCTACACACAGTCAATAACCCCTTTTAGAAGCGGAATTTATTTATGAG 472	
DB	448 GGGAGACTCATCTACACACAGTCAATAACCCCTTTTAGAAGCGGAATTTATTTATGAG 507	
QY	473 ATGGAACAATTTGGAATCCACAATGGTGTGTAATCTCTGGAATCTTGAATGGCTGACATGTC 532	
DB	508 ATGGAACAATTTGGAATCCACAATGGTGTGTAATCTCTGGAATCTTGAATGGCTGACATGTC 567	
QY	533 AGACTGAAATAAGATTAAGTCTGCAATTAATCTGAAAAACCTCTGTGTCAATTAACCAAGA 592	
DB	568 AGACTGAAATAAGATTAAGTCTGCAATTTTCTGAGAAACCTCTGTGTTCAATTTACCAAGA 627	
QY	593 AAAAAATTAATAAATCTAGATTTAGGGTGAAGCTGACACTAGAGAGGCTCGAGAGAGAT- 651	
DB	628 AAAAAATTAATAAATCTAGATTTAGGGTGAAGCTGACACTAGAGAGGCTCGAGAGAGATG 687	
QY	652 -----GACGATGATAGGGTATCTCCACTGTACTCCAAAAATGTCCTGCTGCTGAGAGAGAGATG 706	
DB	688 ATGACGACCATGATTAAGGCATCTCCACTGTCTTTCACAGATGTCCATAGGCTGAGAG 747	
QY	707 TATCTCTTAATAAGCGACAATGAGTTTCAAGTCGAGGCATTTACAGCGGAGTGTGGTTATG 766	
DB	748 TATCTCTTAATAAGTCACAATGAGTTTCAAGTCGAGGCATTTACAGCGGAGATGTGGGTATG 807	
QY	767 GCTTGCAGGCTGATCGTTGGACAGAGTACAGCATACAGACGATGAACACAGATTAACCTGG 826	
DB	808 GCTTACAGGCTGACCGCTGGACAGAGTACAGCATACAGACGATGAAGCGGACCAACTTGG 867	
QY	827 AACTAATCTTTGATTTTTTCGAAGAAGATCTCAGTGAGCAGCTAGTTTCAGGGTGATGCC 886	
DB	868 AACTCATCTTTGATTTTTTCGAAGAAGATCTCAGTGAGCAGCTAGTTTCAGGGTGATGTTTC 927	
QY	887 TTCTTGGAACATGTGGGTACAGCTTGTCTTTTATCATCCACCATTTGCTGAGAGTGGAAGA 946	

Db	928	TTCTCTGACATGTGGGCACAGCATGCCTCTGTCACTACCATTTGCTGAGAGTGAGAA	987
Qy	947	GTGCTGGAAATCTTACTCTTTCCCATCATGACGAGAAATTCGGGAAAAACAATAGGCAAG	1006
Db	988	GGCTGGAAATCCCTTCTCTTTCCCATCATGAGCGAGAGTTCCAGAAAAACATTAGGCCAAG	1047
Qy	1007	TGAGAGTTGACTATATAATTTAATTAAGCCATTACAGGATACAGTTGTGACATGAATCTTT	1066
Db	1048	TCAGAGTTGATTTTATCATCATCAAGCCATTACGAGATAGTTGTCTTATGCACTCTT	1107
Qy	1067	CATTTTCCAAGTATTCGAAGCCAGAAATACCATTTGGATGTTGGCCATCGAGGTGCAGAA	1126
Db	1108	CATTCTCAAGTATTTGAAACCAAGAAATACCATTTGGATGTTGGACATCGTGGTGCAGGA	1167
Qy	1127	ACTCTACAACTGCTCCACGCTGGCTAAAGTTTCAAGAAAAATACATTTGCTTCTTTTAAGAA	1186
Db	1168	ACTCAACAACTGCCACGCTGGCTAAAGTACAGAAAAATACATTTGCTTCTTTTAAGAA	1227
Qy	1187	ATGCTGCTAGTCATGTTGTCGACCTTTGTAGAAATTTGACGTACACCTTTCAAAGGACTTTG	1246
Db	1228	ATGCTGCACCATGTTGTCAGCACTTTGTGNAATTTGATGTGCCACTTTCAAAGGACTTTAG	1287
Qy	1247	TGCCGTGGTATATCATGATCTTACCTGTTGTTTGACTATGAAAGAAATTTTGAATGCTG	1306
Db	1288	TGCCGTGAGTGATCATGATCTCACCTGCTGTTTAACATATGAAAGGAAATATGAAGCTG	1347
Qy	1307	ATCCAGTTGAATTTATTTGAAATTTCCAGTAAAGAAATTTAACTTTTGACCAACTCCAGTTGT	1366
Db	1348	ATCCAGTTGAATTTGTTTGAATTTCCAGTAAAGAAATTTAACTTTTGACCAACTCCAGTTAT	1407
Qy	1367	TAAAGCTCACTCATGTGACTGCATGAAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGG	1426
Db	1408	TGAAGCTTTCTCATGTGACTGCATTAATAACCAAGACCAAGAAACAAATGTATGCCTGAGG	1467
Qy	1427	AGGAAATTTCCCTTTTCGAAAAATCAGCCATTTCCCTCTCTTAAGATGGTTTTAGAGCTTT	1486
Db	1468	AGGAAATTTCCCTTTCTGAAACCAACCAATTTCCCTCTCTTAAGATGGTTTTAGAGCTCAT	1527
Qy	1487	TGCCAGAAATGTAGGGTTTTAACTTGAATAAAATGGATCTGCCAGCAAGGATGGAA	1546
Db	1528	TGCCAGAAATGTAGGATTTAATATAGAAATAAATGGATTTGCCAACACAGGGATGGAG	1587
Qy	1547	TGTGGGATGGTAACTTATCAACATATTTTGACATGAATCTGTTTTGGATATATATTTAA	1606
Db	1588	TATGGGACGGCACTTATCGACATATTTTGATATGAATGCATTTTTTGGATATATATTTAA	1647
Qy	1607	AACTGTTTTAGAAATTTCTGGGAAGAGGAGATAGTGTCTTCTCAATTTGATGTCAGATA	1666
Db	1648	AACTGTTTTAGAAATTTCCGGGAAGAGGAGATAGTATTTCTTCTATTTTGATGTCAGACA	1707
Qy	1667	TTTGCACAATGGTTTCGCAAAACAGAAACAAATATCCGATACTATTTTAACTCAAGGAA	1726
Db	1708	TCGTGCAATGGTTTCGCAGAAACAAATATCCCATATTTATTTTGTGACCCAGGAA	1767
Qy	1727	AACTCGAGATTTTATCTGAACTCATGACCTCATGAGTCTCGGACACCCCAATTCGAATGA	1786
Db	1768	AGCTCGACATTTACCTCGAATCTCATGACCTCATGAGTCTCGGACACACCCCAATTCGAATGA	1827
Qy	1787	GCTTTGACAGTTTGAATCTTACTGGGATTAATGTACATCTGAAGACTTGTCTCAGAA	1846
Db	1828	GCTTTGACAGTTTGAATATTTTGGGATTAATGCCATCTACTGAAGATCTCTCTTAGAA	1887
Qy	1847	ACCATCTCTATTTTCAAGAGCCAAAGCTAAGGACTAGTCTATATTTCTGCTGGGGTGATG	1906
Db	1888	ACCATCTCTATGTCGAAGAGCCAAAGATNAGGATTTGGTCAATTTCTGCTGGGGTGATG	1947
Qy	1907	ATACCAATGATCTCTGAAAAACAGAGGAAATTTGAAGAACTTTGGAGTTAATTTGCTCTAATTT	1966
Db	1948	ATACCAATGATCTCTGAAAAACAGAGGAACTGAAGGAAATTTGGAGTTAATTTGCTCTAATTT	2007
Qy	1967	ATGATAGG	1974

Db	2008	ATCATAGG	2015	
	RESULT 15			
	ABA09667			
ID	ABA09667	standard; DNA; 2477 BP.		
XX	AC	ABA09667;		
XX	DT	15-JAN-2002 (first entry)		
XX	DE	Human bone marrow expressed oligonucleotide SEQ ID NO: 176.		
XX	KW	Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnery;		
KW	KW	antiinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;		
KW	KW	antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;		
KW	KW	antitumor; fungicide; antidiabetic; antiaethmatic; antiallergic;		
KW	KW	immunostimulant; analgesic; cerebroprotective; antianaemic; infection;		
KW	KW	nervous system disorder; autoimmune disorder; inflammation; allergy; ds.		
XX	OS	Homo sapiens.		
OS	PN	W0200174836-A1.		
XX	PD	11-OCT-2001.		
XX	PF	30-MAR-2001; 2001WO-US010472.		
XX	PR	31-MAR-2000; 2000US-00540217.		
PR	PR	23-AUG-2000; 2000US-00649167.		
PR	PR	30-NOV-2000; 2000US-0250583P.		
XX	PA	(HYSE-) HYSEQ INC.		
XX	PI	Tang YT, Liu C, Drmanac RT, Ford JE, Boyle BJ;		
XX	DR	WPI; 2001-626375/72.		
XX	PT	New bone marrow-expressed nucleic acids and polypeptides, useful for		
PT	PT	diagnosis, treatment of inflammatory, autoimmune, neurological, cancer		
PT	PT	and increasing hematopoiesis, stem cell survival and bone growth and		
PT	PT	remodeling.		
XX	PS	Claim 1; Page 254-255; 380pp; English.		
XX	CC	The present invention relates to bone marrow expressed polynucleotides		
CC	CC	and proteins. These sequences can be used in the treatment of		
CC	CC	inflammatory conditions (eg arthritis, Crohn's disease), cancer, central		
CC	CC	and peripheral nervous system diseases and neuropathies, such as		
CC	CC	Alzheimer's, Parkinson's and Huntington's diseases, spinal cord		
CC	CC	disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid		
CC	CC	cell disorders, platelet disorders, stem cell disorders, bone		
CC	CC	degenerative disorders, autoimmune disorders, for example multiple		
CC	CC	sclerosis, diabetes and arthritis, viral and bacterial infections,		
CC	CC	allergies and blood coagulation disorders. The present sequence is a DNA		
CC	CC	of the invention		
XX	XX			
SQ	Sequence	2477 BP; 681 A; 500 C; 642 G; 654 T; 0 U; 0 Other;		
	Query Match	43.9%; Score 1406; DB 5; Length 2477;		
	Best Local Similarity	95.7%; Pred. No. 0;		
	Matches 1482; Conservative	0; Mismatches 0; Indels 66; Gaps 1;		
Qy	616	AGGGTGAAGCTGCACATAGAGGCTCGAGGAAGATCAGATGATGGGTATCTCCACT	675	
Db	996	AGGGTGAAGCTGCACATAGAGGCTCGAGGAAGATCAGATGATGGGTATCTCCACT	1055	
Qy	676	GTACTCCAAAATGTCATAGCTTGGAGATATCCTTAATAGCGCAATGAGTTCAG	735	
Db	1056	GTACTCCAAAATGTCATAGCTTGGAGATATCCTTAATAGCGCAATGAGTTCAG	1115	
Qy	736	TGCAGGCATTCCACAGCCGGAGTGTGGTTATGGCTTTCAGAGCTGATCGTTCGACAGAGTAC	795	


```
QY 1876 AAGGACTAGTCATATTCCTGGGGTGATGATACCAATGATCCTGAAACACAGAGGAAA 1935
Db |||||
QY 2190 AAGGACTAGTCATATTCCTGGGGTGATGATACCAATGATCCTGAAACACAGAGGAAA 2249
Db |||||
QY 1936 TTGAAGGAACTTGGAGTTAAATGCTCTAATTTATGATAGGATATATGATTGGATGCTGAA 1995
Db |||||
QY 2250 TTGAAGGAACTTGGAGTTAAATGCTCTAATTTATGATAGGATATATGATTGGATGCTGAA 2309
Db |||||
QY 1996 CAACCAATATATTCAGTGGAGCAATTGGAACGCTGAAGCAGGAAATTGCCAGAGCTT 2055
Db |||||
QY 2310 CAACCAATATATTCAGTGGAGCAATTGGAACGCTGAAGCAGGAAATTGCCAGAGCTT 2369
Db |||||
QY 2056 AAGAGCTGTTTGTGTCCTGCTTTAGCCGCTTTGTTCCCTCATCTTTGTTGTTGGGAGTCT 2115
Db |||||
QY 2370 AAGAGCTGTTTGTGTCCTGCTTTAGCCGCTTTGTTCCCTCATCTTTGTTGTTGGGAGTCT 2429
Db |||||
QY 2116 GATATCCATGTGGATGCCAACGGCAATTGATAACGTGGAGAAATGCTTAG 2163
Db |||||
QY 2430 GATATCCATGTGGATGCCAACGGCAATTGATAACGTGGAGAAATGCTTAG 2477
Db |||||
```

Search completed: July 2, 2005, 13:08:15
Job time : 1645 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 12:40:51 ; Search time 9973 Seconds
(without alignments)

12236.443 Million cell updates/sec

Title: US-10-047-855-4

Perfect score: 3206

Sequence: 1 gtcgaccacgcgtccggc.....aaaaaaagcgccgcg 3206

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_htc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2038	63.6	3142	3	AK030645	AK030645 Mus muscu
2	2032	63.4	3614	3	AK050318	AK050318 Mus muscu
3	1891	58.7	3248	3	BC006887	BC006887 Mus muscu
4	1794.6	56.0	3578	3	AK031658	AK031658 Mus muscu
5	1570.6	49.0	2604	3	AK009563	AK009563 Mus muscu
6	1420.2	44.3	2228	3	AK051728	AK051728 Mus muscu
7	1170.8	36.5	1618	3	AK049491	AK049491 Mus muscu
8	962.6	30.0	998	5	EX344122	EX344122 BX344122
9	921.4	28.7	1006	5	EX344123	EX344123 BX344123
10	846.8	26.4	982	4	BM457674	BM457674 AGENCOURT
11	835	26.0	1128	4	BM478430	BM478430 AGENCOURT
12	816	25.5	973	5	BQ049943	BQ049943 AGENCOURT
13	815.2	25.4	948	5	BUI45581	BUI45581 AGENCOURT
14	809.4	25.2	811	3	CR623938	CR623938 full-leng
15	799.8	24.9	811	1	AL520917	AL520917 AL520917
16	758.6	23.7	818	5	BX403389	BX403389 BX403389
17	749.2	23.4	758	6	CD520772	CD520772 AGENCOURT
18	746.8	23.3	786	1	AL520916	AL520916 PUBMED
19	739.2	23.1	823	1	AL046038	AL046038 DKFZP434B
20	731.4	22.8	754	7	CN303372	CN303372 170006000
21	728.2	22.7	776	4	BG742916	BG742916 602632004
22	722.2	22.5	735	7	CR768972	CR768972 DKFZP468H
23	706.8	22.0	855	5	BQ962080	BQ962080 AGENCOURT
24	699.8	21.8	821	4	BI822275	BI822275 603036411

25	697	21.7	783	6	CD466056	CD466056 LeukoN2.1
26	692.8	21.6	696	2	AW327746	AW327746 dr01g04.x
27	687	21.4	711	5	BX507382	BX507382 DKFZP779K
28	686.4	21.4	688	4	BM786845	BM786845 K-EST0065
29	681	21.2	731	6	CB310051	CB310051 AGENCOURT
30	670.4	20.9	795	4	BG623187	BG623187 602648092
31	658	20.5	672	5	BX091262	BX091262 BX091262
32	653.4	20.4	728	7	CN303370	CN303370 170004243
33	644.2	20.1	746	7	CN303375	CN303375 170005318
34	635.8	19.8	833	4	BI144681	BI144681 602909978
35	633.4	19.8	768	1	AUI136888	AUI136888 AUI136888
36	632	19.7	901	6	CB196769	CB196769 AGENCOURT
37	631.2	19.7	805	6	CA317224	CA317224 UI-M-FW0-
38	627.6	19.6	958	5	BUS03488	BUS03488 AGENCOURT
39	623	19.4	691	7	CK833283	CK833283 4057562 B
40	620	19.3	656	7	CK001576	CK001576 AGENCOURT
41	616	19.2	619	5	BQ002886	BQ002886 UI-H-E11-
42	614.2	19.2	654	4	BI838003	BI838003 603083846
43	608.8	19.0	841	1	AJ820056	AJ820056 AJ820056
44	608.8	19.0	656	4	BI561344	BI561344 603255105
45	603.8	18.8	665	6	CD640091	CD640091 AGENCOURT

ALIGNMENTS

RESULT 1	AK030645	3142 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK030645	3142 bp	mRNA	linear	HTC 03-APR-2004
DEFINITION	Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430401018 product:hypothetical Glycero-phosphoryl diester phosphodiesterase/Glycosyl hydrolase, starch-binding domain containing protein, full insert sequence.				
ACCESSION	AK030645				
VERSION	AK030645.1	GI:26326638			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipette sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
REFERENCE	5				

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3142)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <http://genome.gsc.riken.jp/> URL: <http://fantom.gsc.riken.jp/> Location/Qualifiers

FEATURES source

1..3142

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:5430401018"

/db_xref="taxon:10090"

/clone="5430401018"

/tissue_type="head"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stages="6 days neonate"

157..2184

/note="unnamed protein product: hypothetical Glycophosphoryl diester phosphodiesterase(Glycosyl hydrolase, starch-binding domain containing protein (InterPro|IPR004129, InterPro|IPR02044, evidence: InterPro)

putative"

/codon_start=1

/protein_id="BAC27063.1"

/db_xref="GI:26326639"

/translation="MTPSQVTFEIRGTLLPGEVFAICGSDALGNWNPQNAVALIN ETDGSLWKAVIALNRGVSVKYPFGCFLEPKTGPCQVIVHKWETHLPQSIITPL ESEIILDQQRGHNGVFTLDSGLTCTQTEIRLRLHSEKPPVSIKKPKSRFRVK LTLGLEDEDDDDKVSFTVLHKMSNLEISLSDNFEKFKHRSQPECGYGLQPDRT EYSTQTFMDNLELFDFFEDLSHVVGQDVLFGHVFTACLSTLSABSRAGILT LPIMSRNRKTIKRVYDFIIKPLPGYSCSMQSFYSKYWKPRIPLDYVHRGAGNLT TAKLAKVQENTIALRNAASHGAAPFVDFVHLSKDFVPVYVHDLTCCILTKMRKYEAD VELPFI PVKELTFDQILLKLSHTALATKDRKOSLYBEENFFSENQFPPLKRVLES LPENVGFNIEIKWICQHRDGYDGNLSYFDQNTVFLDILILKTVLENSKRRI VFSFSD ADICTWRQKQNYFIFELTQGSQDIYFELMDLSRTPPIAMSFQENILGINAHTE DLLRNPSYVQEAAGLVFCWDDNDPNDPNDKLVKFGVNGLIYDRIDWNPQPNL FQVEQLRLKQELPBLKNCLEPTVSHFSPSSFCVEPDIHVDANGIDSVENA"

3125..3130

/note="putative"

3142

/note="putative"

polya_signal

polya_site

ORIGIN

Query Match 63.6%; Score 2038; DB 3; Length 3142;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 2579; Conservative 0; Mismatches 450; Indels 81; Gaps 19;
QY 86 GACGCGGACCTGCGGACTAGCGAACCCGAGACGACATCAATAAATAATCATCAGAA 145
DB 98 GCCGCGCCGAGGCTCTCGGCGCGGCCGCGAGCGGAGAAATCCTAAAAATTCATCAGAA 157
QY 146 TGACACCTTCTCAGGTGCTTTTGAATAAGAGAACTCTTTTACCAGAGAGATTTTTCG 205
DB 158 TGACACCTTCTCAGGTGCTTTTGAATAAGAGAACTCTTTTACCAGAGAGATTTTTCG 217
QY 206 CGATATGTGGAAGCTGTGATGCTTTGGGAACTGGGAATCTCAAAATGCTGTGGCTCTTC 265
DB 218 CAATATGTGGAAGCTGTGATGCCCTGGGAACCTGGAAATCTCAAAATGCTGTGGCTCTTA 277
QY 266 TTCCAGAGAAATGACACAGAGTGAAGCATGCTATGGAAGCAACCATTTGTACTCAGTAGAG 325
DB 278 TTAATGAAACGAGACAGGAGACAGTGTGTTGTGGAAGCAGTGATTGCTCTCAATAGAG 337
QY 326 GAGTATCAGTTCAGTATCGCTACTTCAAGGGTACTTTTGTAGAACCAAGACATATCGGTG 385
DB 338 GAGTGTCAAGTGAAGTACCGCTACTTCAGAGGCTGCTTTTGTAGAACCAAGACATATCGGTG 397
QY 386 GTCATGTCAAGTGAAGTATGTTTCAAGTGGGAGACTCATCTACAAACCAAGCATCAATAACCC 445
DB 398 GTCATGTCAAGTGAAGTATGTTTCAAGTGGGAGACTCATCTACAAACCAAGCATCAATAACCC 457
QY 446 CTTTAAAAACGGAATTAATTTAGCATGGAACAAATTTGGAATCCCAATGGTGTGAAA 505
DB 458 CTTTAAAAAGTGAATCAATTTTACGATGGAACAGTTTGGCATCCCAATGGTGTGAAA 517
QY 506 CTCGTGATCTGATGGCTGACATGTCAGACTGAAATAAGATTAAGTTTGCATTTATCTG 565
DB 518 CACTGGATTCTGATGGCTTACATGTACAGTCAAAATAAGATTTGCGTCTGCAATTTTCTG 577
QY 566 AAAAACTCTGTGTCATTAACCAAGAAATAATAAAATCTAGATTTAGGCGTGAAGC 625
DB 578 AGAAACCTCTGTGTTCAATTTAGCAAGAAATAAGTTTCAAAATAAGATTTAGGCGTGAAGC 637
QY 626 TGACACTAGAAAGCTGGAGGAAAGATGACGAT-----GATAGGGTATCTCCCACTG 676
DB 638 TCACACTCGAGGGTCTGGAGGAAAGATGAAGATGATGATGACGATAAGTCTCTCCCACTG 697
QY 677 TACTCCACAAAATGTCCAATAGCTTGGAGATATCTTTAATAAGGCAATAGATTTCAAGT 736
DB 698 TTCTTCACAAAATGTCCAAGCTGAGATATCTTTAATAAGTGAATAGTGAAGTTCAGT 757
QY 737 GCAGGCATTCACAGCGGAGTGTGGTTATGGCTTGACGCTGATCGTTGGACAGAGATACA 796
DB 758 GCAGGCACTCACAGCCAGAAATGTGGGTATGGCTTTACAGCCCGATCGTTGGACAGAGATACA 817
QY 797 GCATACAGCATGGAAACAGATAAACCCTGGAACCTAATCTTTTGTATTTTTCGAAAGAGATC 856
DB 818 GCATACAGCAATGGAAACAGATAAATCTGGAGCTCATCTTTTGACTTTTGTAGGAAGATC 877
QY 857 TCAGTGAGCAGTAGTTTCAAGGGTATGCCCTTCTCTGGACATGTGGGTACAGCTTGTCTCT 916
DB 878 TCAGTGAGCATGTAGTTTCAAGGGTATGTTCTTCTCTGGACATGTGGGTACAGCATGCTCC 937
QY 917 TATCATCCACATTCGTGAGAGTGGAAAGAGTGTCTGAAATCTTACTCTTCCCATCATGA 976
DB 938 TGTCTTCTACCATTCGTGAGAGTGGAAAGAGCGCTGGAATCCTTACTCTTCCCATCATGA 997
QY 977 GCAGAAATTTCCGAAACCAATAGGCAAGTGTAGAGTTGACTATATATTAATTAAGCCAT 1036
DB 998 GCAGAAATTTCCAGAAACCAATAGGCAAGTGTAGAGTTGACTATATTAATTAAGCCAT 1057
QY 1037 TACCAGGATACAGTTGTGACATGAAATCTTTCAATTTTCCAAAGTATTTGAAGCCCAAGATAC 1096
DB 1058 TACCTGGATATAGTTGTTCTTATGCGAGTCTTCAITTTTCCAAAGTATTTGAACCAAGATAC 1117
QY 1097 CATTGGATGTTGGCCATCGAGGGTGCAGGAAACTCTCAACAACATGCCAGCTGGCTAAG 1156

1118	Db		CAU	TTGGACGTTGGACATCGTGGTGACGGGAAC	CTCAACAGCATGCCAACCTAGCTAAAG	1177
1157	Qy		TTCA	AGAAAAATAC	TATTTGCTTCTTTAAGAAATGCTGCTAGCTCATGGTGACGCTTTGTTAG	1216
1178	Db		TACAG	AAAAATAC	TATCGCTTCTTTAAGAAATGCTCCAGTCAATGGCGCAGCAATTTGTAG	1237
1217	Qy		AA	TTGACGTACACCTTTCAAAGACATTTGTGGCCGTGGTGATATCATGATCTTTACCTGTT		1276
1238	Db		AA	TTGATGCCACTTTCAAAGACATTTGTGCCCGTGGTGATCATGACCTCACCTGCT		1297
1277	Qy		GTTT	GACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTTATTGAAATTTCCAGTAA		1336
1298	Db		GTC	GTACCATGAAGAGAAATATGAAGCTGATCCAGTTGAATTTGTTTGAATATCCCACTAA		1357
1337	Qy		AAGA	TTTAAATTTGACCAATCCAGTTGTTTAAAGCTCACTCATGTGACGTGCACTGAAAT		1396
1358	Db		AAGA	TTTAAATTTGACCAACTCAAGTTATGAAGCTTCTCATGTGACGTGCAATTAATAA		1417
1397	Qy		CTA	AGGATCGGAAGAAATCTGTGGTTTCAGGAGGAAATTTCTTTTCAGAAAAATCAGCCAT		1456
1418	Db		CCAA	GACCGGAACAATCTTTGTTAGAGGAGGAAATTTCTTTCTGAAAAATCAGCCAT		1477
1457	Qy		TTCT	TTCTTTAGATGGTTTATAGTCTTTTGCCAGAAGATGTAGGGTTTAAATTTGAAA		1516
1478	Db		TTCT	TTCTCTTAAAGATGGTTTATGAATCAATGCCAGAAAAATGTAGGATTTTAATATAGAA		1537
1517	Qy		TAAA	TGGATCTGCCAGCAAAAGGATGGAATGTGGGATGGTAACTTATCAACATATTTTG		1576
1538	Db		TAAA	TGGATTTGCCAACACAGGATGGAGTATGGGATGGCACTTATCAACATATTTTG		1597
1577	Qy		ACAT	GAATCTGTTTTTGGATATATTTTAAAAAATCTGTTTTAGAAAAATTTCTGGGAAGGA		1636
1598	Db		ATAT	GAATGTGTTTTTGGATATATTTTAAAAAATCTGTTTTAGAAAAATTTCTGGGAAGGA		1657
1637	Qy		GAAT	AGTGGTTTTCTTCAATTTGATGCAATATTTTGCACAATGTGGTTGGCAAAACGACAA		1696
1658	Db		GAAT	AGTGGTTTTCTTCTTTTATGATGCAGATTTTGTGTA CAATGGTTGGCGCAGAGCAACA		1717
1697	Qy		AATAT	CCGATCATATTTTAACTCAAGAAAAATCTGAGATTTATCTCTGAATCATGCAAC		1756
1718	Db		AATAT	CCCATATATTTTTTGGACCAAGAAAGTCTGATATTTACCCGAACTCATGCAAC		1777
1757	Qy		TCAG	ATCTCGGACAAACCCCAATTCGAATGAGCTTTGCAAGTTTGAATAATCTACTGGGA		1816
1778	Db		TCAG	ATCTCGGACAAACCCCAATTCGAATGAGTTTGGACAGTTTGAATAATTTTGGGA		1837
1817	Qy		TAAAT	GTACATCTGAAGACTTGCTCAGAAACCCATCCTATATTTCAAGAGGCAAAAGCTTA		1876
1838	Db		TAAAT	GCCCATACTGAAGACCTCTTGAAGAACCCATCCTATGTGTCGAAGGCAAAAGCTTA		1897
1877	Qy		AGGAC	TAGTCAATATCTGCTGGGGTGCATGATACCAATGATCTCTGAAAAACAGAAAGAAAT		1936
1898	Db		AGGAT	TGGTCAATATCTGCTGGGGTGCATGATACCAACGATCTCTGAAAAACAGAAAGAAAC		1957
1937	Qy		TGA	AGGAACTTGGAGTTAAATGGTCTAAATTTATGATAGGATATATGATTTGATGCTCGAAC		1996
1958	Db		TGA	AGGAACTTGGAGTTAAATGGTCTAAATATATGATAGGATATATGATTTGATGCTCGAAC		2017
1997	Qy		AAC	CAAAATATATTCGAAGTGAGCAATTTGGAAACGCTGGAAGCAGGAATTTGCCAGAGCTTA		2056
2018	Db		AGCC	CAAAATATATTCGAAGTGAGCAGTTGGAGCGCTTGAAGCAGGAATTTGCCAGAGCTTA		2077
2057	Qy		AGAG	CTGTTGTGCCACTGTTAGCCGCTTTGTTCCCTCATCTTTCTGTGGGGAGTCTG		2116
2078	Db		AGAA	CTGTTGTGCCACTGTTAGCCACTTCATCCCTTCTTTCTGTGTGAGGCTG		2137
2117	Qy		ATAT	TCCA	TGTGGATGCCAAACGGCAATTCATTAACGTGGAGAAATGCTTAGTTTTATTGCA	2176
2138	Db		ATAT	TCCA	TGTGGATGCCAAACGGCAATTCATTAACGTGGAGAAACGCTTAGTTTACAGTGC	2197
2177	Qy		GAG	GTCA	ATTTTGGGGCGTGCAACGCTGTGTTCTGGGTATTCATTTTTCATCATCTGAGCAAT	2236

Db	2198	GAGGACATTTAGGGGACACCGCTGCCCTCATG-----TTTAGTTTTTCCACATCGAGCACT	2255
Qy	2237	GTTGATCTATGCGCTTTTGGCGCTTCTCAGTTTCAATGAAGCAATAATGAAGTATTTAACTCT	2296
Db	2254	GCTG----TTGGCTCTTAGGTTTTCAGTCCCAATGAAGCAATAATGAAGTATTTTACTAT	2309
Qy	2297	TTCACTACAGTTCTTGCAAGTATGCTA-----TTTAAATATCTTGGCCAGGTATATAATT	2349
Db	2310	TTCATTACAGGTCCTGCGCAAGATATCAAGTATGCTATTGTCTCACTTCTGCCAGGTATATAATT	2369
Qy	2350	GCCAGTCAGTCTCTTTATAGTGTGAAATAATTTATTTGGTTAGTAAATATAAATATTTTAAACT	2409
Db	2370	ACCAATCAGTCTCTGTACAATAAGAAATATACTG-----TTTAAATCTTCAGTG	2419
Qy	2410	AAATATATAAATCTATAATGTTTAAACATATGTTCAATTTAAAGCATACGACTTTTGAAATTA	2469
Db	2420	AGATATCTAAGCCCTATAAATGGTGA---ATCTTCATTTAAAG-----TTGGAATTA	2467
Qy	2470	ACTATATAAATAGCTCATATTTTACATTCAGAGCTTTTTCATTTGATCAGGCTCTGAAATCTT	2529
Db	2468	ACGATATAAATATCTCATCTATAAGTTCAAGGCTTTTTCATTTGATTAAGTCTTTAAATCTT	2527
Qy	2530	TAGCACTTTAAGGAAAAATGATATGTCATAATTTATACCTGACCATCGAAAAAATAAGTACCT	2589
Db	2528	CAGTGCTTGAAAAATGA---TTGCATAATTTATACCTGACCAATGGAATACTAAGTACCT	2585
Qy	2590	CAATGCATGCATTTGCACTGGTGAATTTCCAACTGCACAAATCTTTTGCCACTTTGTATA	2649
Db	2586	C-AATGCATGCATTTTGCACTGGTGGCTCCAGCTGCA CAGTCTGTGT-CATCCATGTACA	2643
Qy	2650	TAGTATTTTTTACATGGTTTGACATGCACACACACATTTTCATTTCAGTATGAACCTT	2709
Db	2644	TAGGTATCTTTT--ACATGGTTGATAGGAACAC--GCACCATTTTTCCTTTTAGTATAAACCC-T	2699
Qy	2710	GAGGCTCTGCATTTTTCACCTTTTAAACCAACAGCCCTGAAGGTGAACCTTCGAACTTTGT	2769
Db	2700	CAGATTCTGCATTTCTCACTGACCCAGCCAGCTGCAGATGAACCTTCAAACTTGT	2759
Qy	2770	TTCAATAATCTTTCAAAGTTGTTTTTACATCAATGTGTTAAAAATTTCAAATGCTGCAGGGT	2859
Db	2760	CTCATAAACCATTTCAAA-----GATATTTAAAGTTCCAGAAATGCTGCAGGGT	2805
Qy	2830	AATTTTATGTATAAATAATTTAGTAAAGAAAGTATGTATTTGCATACCTTAGTAGNAATAGAT	2889
Db	2806	AACTTAATGTATAAAGTATTTGTAAG---AAGTATATATTTGCATATATAGTAGTGTAGAT	2862
Qy	2890	CACAACATACAAATTTCAATTTCAGTGCATGCTTTTAGGTGTTAAAGCATCAGATTTGACATGT	2949
Db	2863	CAGAAGGTATCAATTTGACATCAATGCATGCTTTTAGGTTTTTAAGCATGAGATTTGTACATGT	2922
Qy	2950	TTACTGTTAGGTCTTCGCATCTGTGGTGTCTAGGTGAGTATGAGAAAGATGCAAGGACTGG	3009
Db	2923	TTACTGTTAAGTCTCTGCATC--TGGTGTAGGTGAG-----GGAGATGTTTAAGGACATA	2975
Qy	3010	ACGTATTTTGTGTGCCTAAAAAAGGCTGTTTGTAGGCGTTTAAATATGCTTATTTT	3069
Db	2976	AAATATTTTGTGCGCTTAAAAAACAAGTCTGTTT--TAGACACTTTGAGTATGCTTATTTA	3034
Qy	3070	GTGTGCTCTCTCACTACCTATTTACACACTGTGTCTTGTGGGTTTGTGTTTGTATGTGCGTG	3129
Db	3035	GTGTCTCACTACCTTAGTAAACATAAGTGTCTCACTACCTAGTAACAT--AATTTGCTGTG	3092
Qy	3130	TGTTATACAGTAGTATAATTTCCATGAGAAAAATAAATGATCTCGAATTC	3179
Db	3093	TGTTGCATAGTAGTATAATTTCCATGAGAAAAATAAATGATCTCGAATTC	3142

RESULT 2
AK050318
LOCUS
DEFINITION

AK050318 3614 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male liver tumor cDNA, RIKEN full-length
enriched library, clone: C730037804 product: hypothetical
glycerophosphoryl diester phosphodiesterase/Glycosyl hydrolase,

starch-binding domain containing protein, full insert sequence.
AK050318
VERSION AK050318.1 GI:26093898
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 3614)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissue was provided by William A. Held, Roswell Park Cancer

Institute, Department of Molecular and Cellular Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES
Location/Qualifiers
source
1..3614
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C730037B04"
/db_xref="taxon:10090"
/clone="C730037B04"
/sex="male"
/cell_type="tumor"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
117..2145
/notes="hypothetical Glycerophosphoryl diester phosphodiesterase/Glycosyl hydrolase, starch-binding domain containing protein (InterPro|IPR004129, InterPro|IPR002044, evidence: InterPro)
putative"
3596..3601
/notes="putative"
3614
/notes="putative"
polyA_signal
polyA_site

ORIGIN
Query Match 63.4%; Score 2032; DB 3; Length 3614;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 2584; Conservative 0; Mismatches 450; Indels 82; Gaps 20;
QY 86 GAGCGGACCTGCGGACTAGCGAACCGGAGCAGCATCATATAATAATAATCCATCAGAA 145
DB 58 GCCGCGCCGAGGCTCTCGGCGCGCGGAGCGGCGAGAAATCTTAATTCATCAGAA 117
QY 146 TGACACCTTCTCAGGTTGCTTTGAATAAGAGGAACTCTTTTACCAGGAGAGGTTCTTG 205
DB 118 TGACACCTTCTCAGGTCACCTTTGAATAAGAGGAACTCTTTTACCAGGAGAGGTTCTTG 177
QY 206 CGATATGTGAAGCTGTGATGCTTTGGAACTGGAACTCTCAAAATGCTGTGGCTCTTC 265
DB 178 CAATATGTGAAGCTGTGATGCCCTGGAACTGGAACTCTCAAAATGCTGTGGCTCTTA 237
QY 266 TTCAGAGATGACACAGGTGAAGCATGCTATGGAAAGCAACCATTTGTTACTCAGTAGAG 325
DB 238 TTAATGAAACGACACAGGAGACAGTGCTGTGGAAAGCAGTGATTCCTCAATAGAG 297
QY 326 GAGTATCAGTTTCAGTATCGCTACTTCAAAGGGTACTTTTAGAACCAAGACTATCGGTG 385
DB 298 GAGTTCAGTGAAGTACCGCTACTTTCAGAGGCTGCTTTTAGAACCAAGACTATCGGTG 357
QY 386 GTCATGTCGAAGTGAAGTTCACAAAGTGGAGACTCAATCTACAAACCGATCAATAACC- 444
DB 358 GTCCATGTCAAGTCATAGTTTACAAAGTGGAGACTCATCTACAAACCGATCAATAACCC 417
QY 445 CTTTAAAGGCGAAATATTTATGACATGGACAAATTTGGATCCCAATGGTTGAA 504
DB 418 CTTTAAAGAGTGAATATCATTTTGCATGGAGTTCGATCCCAATGGTTGAA 477
QY 505 ACTCTGATTTCTGATGGCTGACATGTCAGATGAAATAAGATTACGTTTGCATTATCT 564
DB 478 ACATGGATTTCTGATGGCTTACATGTCAGATGAAATAAGATTGCGTTCGATTTCT 537
QY 565 GAAAAACCTCTGTGTCAATAACCAAGAAAAAATTAATAAATCTAGATTAGGGTGAAG 624
DB 538 GAGAAACCTCTGTGTTCAATTAGCAAGAAAAAGTTCAAAAAATCGAGATTAGGGTAAAG 597
QY 625 CTGACACTAGAGGCGCTGGAGGAAGATGACGAT-----GATAGGGTATCTCCCACT 675

[illegible]

```
QY 2829 TAATTAATGTAATAATATTAGTAAGAAAGATGATGATTCATCTTGTAGTAGATAGA 2888
Db 2766 TAACCTTAATGTAATAAGTAATTTGTAAG---AAGTATATATGTCATATATAGTAGTGAGA 2822
QY 2889 TCACAAATACAAATTCATTAATCAGTGAATGCTTTAGGTGTTAAAGCATGAGATGTCATAG 2948
Db 2823 TCAGAAAGTATCAATTTGACTCAATGCATGCTTTAGGTGTTAAAGCATGAGATGTCATAG 2882
QY 2949 TTTACTGTTAGGTCCTTGTCATCTGTGGTGTAGTGTAGTATGAGAAAGATGTCAGGACTG 3008
Db 2883 TTTACTGTTAAGTCTTGTCATC---TGGTGCTAGGTGAG-----GGAGATGTTAAGGACAT 2935
QY 3009 GACGTATTTTGGCTCAAAAAAAGCGCTTTGTAGGCTTTTAAATATGCTTAATTT 3068
Db 2936 AAAATATTTTGGCTCAAAAAAAGCGCTGTTT-TAGACACTTTTGGATGCTTAATTT 2994
QY 3069 TGTGTGTCCTCACCTACCTATTACACACTGTTGCTTTGTGGGTTGTTTGTATGTCGCT 3128
Db 2995 AGTGTCTCACTACCTAGTAACATATAGTGTCTACTACCTAGTAACATATGTCG--T 3052
QY 3129 GTGTTATACAGTAGTAAATTTCCATGCGAGAAAAATAAATCTCTGTAATTCCTCAA 3184
Db 3053 GTGTTGCATAGTAGTAATTTCCATGCGAGAAAAATAAATGCTGCTGAATTCCTCAA 3108

RESULT 3
BC006887
LOCUS BC006887
DEFINITION Mus musculus RIKEN CDNA 2310032D16 gene, mRNA (CDNA clone
IMAGE:3596618), containing frame-shift errors.
ACCESSION BC006887
KEYWORDS HTc.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, K.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) National Institutes of Health, Mammalian
MEDLINE Gene Collection (MGC), Cancer Genomics Office, National Cancer
PUBMED Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
2001
REFERENCE
AUTHORS NIH-MGC Project URL: http://mgc.nci.nih.gov
TITLE Contact: MGC help desk
JOURNAL Email: cgapbs-r@mail.nih.gov
REMARK Tissue Procurement: Jeffrey Green M.D.
COMMENT cDNA Library Preparation: Life Technologies, Inc.
```

```
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleaged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 16 Row: 1 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: frame shifted.
FEATURES
Location/Qualifiers
1..3248
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3596618"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal Carcinoma, 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
ORIGIN
Query Match 58.7%; Score 1881; DB 3; Length 3248;
Best Local Similarity 80.5%; Pred. No. 0; Indels 157; Gaps 20;
Matches 2508; Conservative
QY 86 GAGCGGACCTGCGGACTAGCGAAACCGGAGCAGCAGATCAATAAAATAATCCATCAGAA 145
Db 68 GCGCGGCCGAGGCTCTCGCGCGCGCGGAGCGGAGGAGAAATCCTAAATTCATCAGAA 127
QY 146 TGACACCTTCTCAGGTGCTTTGAAATAGAGAACTCTTTTACCAGGAGAAAGTTTTC 205
Db 128 TGACACCTTCTCAGGTGCTTTGAAATAGAGAACTCTTTTACCAGGAGAGGTCCTTG 187
QY 206 CGATATGTGAAGCTGTGATGCTTTGGGAACCTGGAATCTCTCAAAATGCTGTGCTCTTC 265
Db 188 CAATATGTGGAAGCTGTGATGCTGCGGAACTGGAATCTCTCAAAATGCTGTGCTCTTA 247
QY 266 TTCCAGAGATGACACAGGTGAAAGCATGCTATGGAAGCAACCATTTGTACTCAGTAGAG 325
Db 248 TTAATGAAACGAGACAGGAGACAGTGTGTTGGAAGCAGTGATTTCTCTCAATAGAG 307
QY 326 GAGTATCAGTTTCAGTATCGCTTCTCAAGGGTACTTTTGTAGAACCAAGACTATCGGTG 385
Db 308 GAGTGTGAGTACGAGTACCGTCTTTCAGAGGCTGCTTTTGTAGAACCAA----- 355
QY 386 GTCCATGTCAAGTATAGTTTACAAAGTGGAGACTCATCTACACACGATCAATAACCC 445
Db 356 ----- 355
QY 446 CTTTAGAAGCGGAATTTATTGACGATGACAAATTTGGAATCCCAATGGTGTGAAA 505
Db 356 ----AGAAAGTGAATCAATTATTGACATGGACACAGTTTGGCATCCACATGTTGTTGAAA 411
QY 506 CTTCTGATTTGAGTGGCTGACATGTCAGACTGTAATTAAGATTACGTTTCATTAATTCG 565
Db 412 CACTGATTTGAGTGGCTTACATGTCAGACTGAAATTAAGATTGCGTTCGCTTCAATTTCTG 471
QY 566 AAAAACTCTCTGTCTCAATACCAAGAAAAAATTTAAAAAATCTAGATTAGGTTAGGTAAGC 625
Db 472 AGAAACCTCTCTGTTTCAATTAGCAAGAAAAAGTTCAAAAAATCGAGATTTAGGGTAAAGC 531
QY 626 TGACACTAGAGGCTCGGAGAACATGACGAT-----GATAGGTTATCTCCCACTG 676
```



```
QY 2830 AATTATGTTATAAATATTAGTAAAGAAAGATGTTATTCATCTAGTACATAGTATGAT 2889
Db 2700 AACTTAAGTATAAAGTATTGTTAG-----AAGTATATATTCATATATATAGTGTAGAT 2756
QY 2890 CACAACATACAAATTCATTCAGTGCATGCTTTAGGTGTTAAAGCATGAGATTGTACATGT 2949
Db 2757 CAGAAGGTATCAATTTGACTCAATGATGCTTTAGGTTTAAAGCATGAGATTGTACATGT 2816
QY 2950 TTACTGTTAGTCCCTTGCATCTGCTGCTAGTGTGATGATGAGAAAGATGTCGAAGACTGG 3009
Db 2817 TTACTGTTAAGTCCCTTGCATC--TGGTGCTAGGTGAG-----GGAGATGTTTAAAGCAT 2869
QY 3010 ACGTATTTTGTGCTTAAATAAAGAAAGCTGTTTCTAGGCTTTTAAATATGCTTATTTT 3069
Db 2870 AAATATTTTGTGCTTAAATAAAGAAAGCTGTTT--TAGACACTTTTGATGATGCTTATTTA 2928
QY 3070 GTGTCTCTCTACTACCTATTACACACTGTTGCTTTGGGTTGTTTGTATGTGGGTG 3129
Db 2929 GTGTCTCTACTACCTAGTAAACATAGTCTCTACTACCTAGTAAACAT--AATTGCTGTG 2986
QY 3130 TGTATACAGTAGTAAATTTCCATGCGAGAAATAAATGCTGGAATTCCTCAAA 3184
Db 2987 TGTTCAGTAGTAGTAAATTTCCATGCGAGAAATAAATGCTGGAATTCCTCAAA 3041

RESULT 4
AK031658
LOCUS
DEFINITION
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030473012 product:hypothetical
Glycophosphoryl diester phosphodiesterase/Glycosyl hydrolase,
starch-binding domain containing protein, full insert sequence.
AK031658
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
2
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
3
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861
4
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 695-690 (2001)
MEDLINE
11076861
5
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
```

```
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
6 (bases 1 to 3578)
JOURNAL
REFERENCE
AUTHORS
Adachi, Y., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Iehii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
FEATURES
source
Location/Qualifiers
1. 3578
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6030473012"
/db_xref="taxon:10090"
/clone="6030473012"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
misc_feature
205. 2109
/notes="hypothetical Glycophosphoryl diester
phosphodiesterase/Glycosyl hydrolase, starch-binding
domain containing protein (InterPro|IPR004129,
InterPro|IPR002044, evidence: InterPro)
putative"
polyA_signal
3560. 3565
/note="putative"
polyA_site
3578
/note="putative"
ORIGIN
Query Match 56.0%; Score 1794.6; DB 3; Length 3578;
Best Local Similarity 79.3%; Pred. No. 0;
Matches 2465; Conservative 0; Mismatches 454; Indels 188; Gaps 21;
QY 86 GACCGGACCTGGGACTAGGAAACCGGAGCAGCATATATAAATAAATATCCATCAGAA 145
Db 146 GCCCGCGCCAGGCTCTCGCGCGCGCGCGGAGCGGAGAAATCTAAATTCATCAGAA 205
QY 146 TGACACCTTCTCAGTTGCTTTGAAATAAGAGAACTCTTTTACCAGGAGAGTTTGTG 205
Db 206 TGACACCTTCTCAGTTGCTTTGAAATAAGAGAACTCTTTTACCAGGAGAGTTTGTG 265
QY 206 CGATATGTGGAAGCTGTGATGCTTTGGGAAA--CTGGAATCTCAAAATGCTGGCTCTT 264
Db 266 CAATATGTGGAAGCTGTGATGCTTCCTGGAAACCTGGAAATCTCAAAATGCTGGCTCTT 325
```

Qy	265	CTTCCAGAGATGACACAGGTTGAAAGCAATGCTATATGAAAGCAACCAATTTGTAATCTCAGTAGA	324
Db	326	ATTTAATGAAAACGAGACAGGAGACAGTGTGTGTGTGGAAGCAGTGATTTGCTCTCAATAGA	385
Qy	325	GGAGTATCCAGTTTCAGTATCGCTACTTCCAAAGGGTACTTTTTATAGAACCAAGACATATCGGT	384
Db	386	GGAGTGTCTCAGTGAAGTACCGCTACTTTCAAGGGTGTCTTTTATAGAACCAAGACATATCGGT	445
Qy	385	GGTCCATGTCAAGTGATAGTTTCACAAGTGGGAGACTCATCTCAACCAACGATCAATAACC	444
Db	446	GGTCCATGTCAAGTCATAGTTTCACAAGTGGGAGACTCATCTCAACCAACGATCAATAACC	505
Qy	445	CCTTTAGAAACGGAAAAATTTATTATGACGATGGACAAATTTGGAATCCCAATGTGGTTGAA	504
Db	506	CCTTTAGAAAGTGAATCAITTTATTGACGATGGACAGTTTGGCATCCACA	554
Qy	505	ACTCTGGATTCTGGATGSGCTGACATGTGCAGACTGAAATAAGATTACGTTTGGCATTTATCT	564
Db	555	-----	555
Qy	565	GAAAAAACCTCTGTGTCAATAACCAAGAAAAAATTAATAAAAAATCTAGATTTTAGGGTGAAG	624
Db	555	-----GGTAAAGCTCACACTCGAGGGT---	576
Qy	625	CTGACACTAGAAGCGCTGGAGGAGATGACGATGATAGGGTAATCTCCACATGTACTCCAC	684
Db	577	-----CTGGAGGAAGATGAAGATGATGACGATAAAGTCTCTCCACATGTCTTCTCAC	630
Qy	685	AAAATGTCANTAGCTTGGAGATATCCTTAATRAGCGACATGATGATTCNAAGTGCAGGCAT	744
Db	631	AAAATGTCACACGCTGGAGATATCCTTAATAGTGAACAATGATGATTCNAAGTGCAGGCAC	690
Qy	745	TCACAGCCGAGTGTGGTTATGCGCTGTGCAGCGCTGATCGTTGGACAGAGTACAGCATACAG	804
Db	691	TCACAGCCAGAAATGTGGTATGCTGTACAGCCCGATCGTTGGACAGAGTACAGCATACAG	750
Qy	805	ACGATGGAAACAGATAACCTCGAACTAATCTTTTGATTTTTTCGAAAGAGATCTCAGTGAG	864
Db	751	ACAAATGGAACCAAGATAATCTGGAGCTCATCTTTTGACTTTTTTGAGGAAGATCTCAGTGAG	810
Qy	865	CACGTAGTTCAAGGTGATGCCCTTCTTCGGACATGTGGGTACAGCTGTCTCTTTATCATCC	924
Db	811	CATGTAGTTCAAGGTGATGTTCTTCCTCGGACACGTGGGCACAGCATGCTCCTGCTCTTCT	870
Qy	925	ACCAATGCTGAGAGTGGAAAGAGTGTGGAAATTTCTTACTCTTCCTCCATCATGAGCAGAAAT	984
Db	871	ACCAATGCTGAGAGTGGNAAGCGCTGGAAATCTTACTCTTCCTCCATCATGAGCAGAAAT	930
Qy	985	TCCCGGAAAAACAATAGGCAAAAGTCAGAGTTGACATATATAATTTAAGCCATTAACGAGA	1044
Db	931	TCCAGAAAAACTATAGGCAAAAGTCAGAGTTGATTTTATCATCATCAAGCCATTAACCTGGA	990
Qy	1045	TACAGTTGTGACATGAAATCTTTCATTTTCCAAAGTATTTGGAAAGCCAGAAATACCATTTGAT	1104
Db	991	TATAGTTGTTCATATGTCAGTCTTTCATTTTCCAAAGTATTTGGAAACCAAGAAATACCAATGGAC	1050
Qy	1105	GTTGCGCAATCGAGTGCAGGAAAACCTCTCAACAACTGCCCCAGCTGGCTTAAAGTTCAAGAA	1164
Db	1051	GTTGGAACATCTGTGTGCAGGGAACCTCAACAACGACTGCCAAGCTAGCTTAAGATACAGAA	1110
Qy	1165	AATACTATGTCTCTTTAAGAAAATGCTGCTAGTCAATGGTCAGCGCTTTGTAGAAATTTGAC	1224
Db	1111	AATACTATCGCTTCTTTAAGAAAATGCTGCCAGTCAATGGCGCAGCATTTGTAGAAATTTGAT	1170
Qy	1225	GTACACCTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACTGTTGTTTGA	1284
Db	1171	GTCCACCTTTCAAAGGACTTTGTGCCCGTGGTATCATGACCTTCACTGCTGCTGCTGAC	1230
Qy	1285	ATGAAAAGAAATTTGATGCTGATCCAGTTTCAATTTATTGGAATTTCCAGTAAAGAAATTA	1344
Db	1231	ATGAAGAGAAATATGAAGCTGATCCAGTTGAAATTTGTTGAAATTTCCAGTAAAGAAATTA	1290
Qy	1345	ACATTTGACCAACTCCAGTTGTTTAAAGCTCACTCATGTGACTGCACTGAAATCTTAAGGAT	1404

1291	ACATTTTGACCAATCCCAAGTTATTGAAGCTTTCTCATGTGACTGCAATAAAAACCAAAGAC	1350
1405	CGGAAGAATCTGTGGTTTCAGGAGGAAAATTCCTTTTCAGAAAAATCAGCCATTTCCCTTCT	1464
1351	CGGAACAATCTTTGTATGAGGAGGAAAATTCCTTTTCGAAAAATCAGCCATTTCCCTTCT	1410
1465	CTTAAGATGGTTTATGAGTCTTTGCGAAGAGATGTAGGCTTTAAACATTTGAAATAAATGG	1524
1411	CTTAAGATGGTTTATGAATCATTCGCAGAAAAATGTAGGATTTAATATAGAAAAATAAATGG	1470
1525	ATCTGCCAGCAAAAGGATGGNAATGTGGGATGGTAACTTATCAACATATTTTTCACATGAAT	1584
1471	ATTTCGCAACACAGGGATGGGAGTATGGGATGGCAACTTATCAACATATTTTTCATATGAAT	1530
1585	CTGTCTTTTGGGATATAATTTTAAAAACCTGTTTATAGAAAAATCTTCGGAAGAGGAGAAATAGTG	1644
1531	GTGTCTTTTGGGATATATATTTTAAAAACCTGTTTATAGAAAAATCTTCGGAAGAGGAGAAATAGTG	1590
1645	TTTTCTTCATTTTGATGACAGATATTTGCACAATGGTTTCGCAAAAGCAGAAACAAATATCCG	1704
1591	TTTTCTCTTTTGATGCAGATATTTGTACAAATGGTTTCGCGAGAAGCAGAAACAAATATCCC	1650
1705	ATACTATTTTAACTCAAGGAAAAATCTGAGATTTATCTTGAACTCATCGAACCTCAGATCT	1764
1651	ATAATATTTTTCAGCCAGGAAAGCTCTGATATTTTACCCGAACTCATCGAACCTCAGATCT	1710
1765	CGGACAAACCCCATTTGCAATGAGCTTTTGCAACAGTTTGAAAAATCTACTCGGGATAAATGTA	1824
1711	CGGACAAACCCCATTTGCAATGAGTTTTCGACAGTTTGGAANAATATTTTGGGATAAATGCC	1770
1825	CATCTGAAAGACTTGTCTCAGAAACCCATCCTCATATTTCAAGAGGCAAAAAGCTAAGGGACTA	1884
1771	CATCTGAAAGACTTGTCTCAGAAACCCATCCTCATATTTCAAGAGGCAAAAAGCTAAGGGATTG	1830
1885	GTCAATATCTGTGGGTGATGATACCAATGATCTTGAAAAACAGAAAGGAAAATTGAGGAA	1944
1831	GTCAATATCTGTGGGTGATGATACCAACGATCTTGAAAAACAGAAAGGAAAATTGAGGAA	1890
1945	CTTGGAGTTAATGGTCTAAATTTATGATGAGATATATGATTTGATGCCCTTGAAACAAACAAAT	2004
1891	TTTGGAGTAATGGTCTAATATATGATGAGATATATGATTTGATGCCCTTGAAACAGCCAAT	1950
2005	ATATTCCAAGTGGAGCAATTTGGAAACGCTGGAAGCAGGAAATTCAGCAGAGCTTTAAGAGCTGT	2064
1951	ATATTCCAAGTGGAGCAATTTGGAGGCTTGAGAGCAAGAAATTCAGCAGAGCTTTAAGAACTGT	2010
2065	TTGTGTCCACATGTTAGCCGCTTTGTTCCTCATCTTTGTGTGGGAGTCTGATATCCAT	2124
2011	TTGTGTCCACATGTTTAGCCACTTCATCCCTCTCTCTTTCTGTGTGGAGCCTGATATCCAC	2070
2125	GTGGATGCCAACGCCATTTGATACGTGGAGAACTTTAGTTTATTTTCACAGAGGTCTAT	2184
2071	GTGGATGCCAACCGCATTTGATAGTGGAGAACGCTTAGTTACAGTGCACAGAGGACAT	2130
2185	TTTGGGGGCGTGACCGCTGTTCTGGGTATTTCACTTTTTTCATCACGTAGCATTTGTGTATCT	2244
2131	TTAGGGGCACACCGCTGCCCTCATG---TTTAGTTTTCACCATCGACCACTGCTG---2182	
2245	ATGCGCTTTTGGGCTTCTCAGTTTCAATGAAGCAATAATGAAGTATTTTAACTCTTTTCACATAC	2304
2183	TTGCGCTTATAGTTTTCAGTCCAAATGAAGCAATATGAAGTATTTTATTTACTATTTTCATTTAC	2242
2305	AGTTCTTTCGAAGTATGCTA-----TTTAAATTAAGTTCGCCAGGTATAATTCGCAGTCA	2357
2243	AGGTCCCGCAAGAAATATCAAGTATGCTATTTGTGCTACCTGTGCAGGTATAATTTACCAATCA	2302
2358	GTCTCTTTATAGTGAGAAAAATTTATTTGGTTAGTAAATATAAAATATTTTAAAACTAAATATAT	2417
2303	GTCTGTGTACAAATAGAAAAATATACG-----TTAAACATTTTCAGTGAGTATCT	2352
2418	AAATCTATAATGTTTAAACATATGTTTCATTTAAAGACATAGCACCTTTGAAATTTAACTATATA	2477


```
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
94..1449
/note="unmanned protein product; hypothetical
Glycerophosphoryl diester phosphodiesterase/Glycosyl
hydrolase, starch-binding domain containing protein
(InterPro|IPR004129, InterPro|IPR002044, evidence:
putative"
/codon_start=1
/protein_id="BAB26361.2"
/db_xref="GI:26365211"
/translation="MEPDNLELIDFDFEEDLSEHVQGVQDLPHGVGTACLSSTIAES
RGAGSLTPIIMSRNKITIGKRVDFIIPKPGVSCSMQSFYKPKPIPLDVGH
RCAGNSTITAKLAKQVENTIASLRNAASHGAFVDFVHLSKDFVYVVDLTCCLTM
KKRYEADPVELPEIPKELTPDQLQLKLSHTALKTDRKQSLYEENPFSENGPFP
SLKVLSELPENGVNIEIKWICQHRDGVWDGNLSYFDNMFVLDIILKTVLENSGR
RIVFSFDADICTMVRQKNKYPILFLTGKSDIYPELMDLSRTPPIAMSPAQENI
SGINAHTEDLLENPSYVQAKAGLVIFCWGDDTNDPENRKLLEFGVNGLIYDRIYD
WMEQPNIFQVQLERLKLQELPELKNCLPTVSHFIPSSFVCEPDIHVDANGIDSVEN
A"
```

ORIGIN

```
Query Match 49.0%; Score 1570.6; DB 3; Length 2604;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 2041; Conservative 0; Mismatches 364; Indels 72; Gaps 18;

QY 715 ATAAGCGCAATGAGTTCGAAGAGATCTCAGTGAGCAGCTAGTTCAGGGTGATGCCCTTCTCTGGA 774
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 775 CCGTATCGTTGACAGAGTACAGCAGTACAGAGCAGTGGACACAGATAACCTGGAACCTAATC 834
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CCGGATCGTTGGACAGAGTACAGCAGTACAGACAAATGGAACAGATATCTGAGGCTCATC 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 835 TTTGATTTTTTCGAAGAAGATCTCAGTGAGCAGCTAGTTCAGGGTGATGCCCTTCTCTGGA 894
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 TTTGACTTTTTTGGAGGAAGATCTCAGTGAGCAGTATGTTTTCAGGGTGATGTTCTTCTCGGA 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 895 CATGCGGTGACAGCTTGTCTTATCATCCAGCATCTGCTGAGAGTGGAAAGAGTGTCTGGA 954
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 CAGTGGGACAGCATGCTCTCTCTTCTACATTCAGTTCAGAGTGGAAAGAGCGCTGGA 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 955 ATTCTTACTCTCCCATCATGACGAGAAATCCCGGAAACAAATAGGCAAGTGCAGAGTT 1014
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 ATCCTTACTCTTCCCATCATGACGAGAAATTCAGAAACAACTATAGGCAAGTGCAGAGTT 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1015 GACTATATAATTAAGCCATTTACCAGGATACAGTGTGTGACATGAATCTTCAATTTTCC 1074
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 GATTTTATCATCATCAAGCCATTTACCTGGATATAGTGTCTATGCAGTCTTCAATTTTCC 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1075 AAGTATGGAAGCAAGAATACATTTGGATTTGGCCATCGAGGTGCAGGAACCTCTACA 1134
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 AAGTATGGAACCAAGAAATACCATTTGGACGTTGGACATCGTGGTGCAGGGAACTCAAC 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1135 ACNACTGCCAGCTGGCTAAAGTTCAAGAAATACATTTGCTTCTTAAAGAAATGCTGCT 1194
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 AGGACTGCCAAGCTAGCTAAAGTACAGGAAATACATTCGCTTCTTAAAGAAATGCTGCC 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1195 AGTCATGGTGCAGCCCTTGTGAGAAATTTGACGTACACCTTTCAAAGGACCTTTGTGCCCGTG 1254
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 AGTCATGGCGCAGCATTTGTAGAAATTTGATGTCCACCTTTCAAAGGACCTTTGTGCCCGTG 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1255 GTATATCATGATCTTACCTGTTGTTGACTATGAAAGAAATTTGATGCTGATCCAGTT 1314
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 GTGATATCATGACCTCACCTGCTGCTGACCATGAAGAGGAAATATGAAGCTGATCCAGTT 600
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1315 GAATTAATTTGAAATTCAGTAAAGAAATTAACATTTTGACCAACTCAGTTGTTAAAGCTC 1374
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 GAATTTGTTTGAATCCAGTAAAGAAATTAACATTTTGACCAACTCAGTTTGAAGCTT 660
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 1375 ACTCATGTGACTGCACTGAAATCTAAGGATCGGAAGAAATCTGTGGTTCAGAGGAGAAAT 1434
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1435 TCCTTTTTCAGAAAATCAGCCATTTCCCTCTCTTAAGATGGTTTATAGTCTTTTGGCCAGAA 1494
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 TTCTTTTTCGAAAATCAGCCATTTCCCTCTCTTAAGATGGTTTATAGATCATTTGCCAGAA 780
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1495 GATGTAGGTTTAAACATTTGAAAATAAATGGAATCTGCCAGCAAAAGGATGGAATGTGGAT 1554
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 781 AATGTAGGATTTAATATAGAAATAAATGGAATTTGCCAACACAGGGATGGAGTATGGAT 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1555 GGTAATCTATCAACATATTTTGAACATGAAATCTGTTTTTGGATATATATTTAAAAACTGTT 1614
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 841 GCGAATCTTATCAACATATTTTGGATATGAAATGTTTTTGGATATATATTTAAAAACTGTT 900
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1615 TTAGAAAATCTCGGAACAGAGAGATAGTGTCTTCTCATTTGATGAGATATTTTGGACAA 1674
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 901 TTAGAAAATCTCGGAACAGAGAGATAGTGTCTTCTTCTTTGATGAGATATTTTGTACA 960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1675 ATGTTTCGGCAAAAGCAGAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAG 1734
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 ATGTTTCGGCAGAGCAGAACAAATATCCCATATATTTTGGACCAAGGAAAGTCTGAT 1020
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1735 ATTTATCTGAATCATGGAACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTGCA 1794
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 ATTTACCCCGAATCATGGAACCTCAGATCTCGGACAAACACCCATTTGCAATGAGTTTGA 1080
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1795 CAGTTTGAATCTACTCGGGATTAATCTACATCTGAAGACTTCTCAGAAACCCATCC 1854
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1081 CAGTTTGAATATTTTCGGGGATTAATGCCATCTGAAGACCTCTTAGAACCCATCC 1140
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1855 TATATTTCAAGAGGCAAAAGCTAAGGGAATAGTTCATATTTCTGCTGGGGTGAATACCAAT 1914
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1141 TATGTCGAAGGGCAAAAGCTAAGGGAATGTCATATTTCTGCTGGGGTGAATACCAAC 1200
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1915 GATCTGAAAAACAGAAAGAAATTTGAAGAACTTGAAGTAAATGTCATATTTATGATAGG 1974
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1201 GATCTGAAAAACAGAAAGAACTTGAAGAAATTTGAAGTAAATGTCATATTTATGATAGG 1260
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1975 ATATATGATTTGGATGCTTGAACCAACCAATATATTTCCAAAGTGGAGCAATTTGAAACGCTG 2034
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1261 ATATATGATTTGGATGCTTGAACCAACCAATATATTTCCAAAGTGGAGCAATTTGAAACGCTG 1320
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2035 AAGCAGGAATTTGCCAGAGCTTAAAGAGCTTTGTGTCCACTTTAGCGCTTTTGTCTCC 2094
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1321 AAGCAAGAAATTTGCCAGAGCTTAAAGAACTTTGTGTCCACTTTAGCCACTTCATCCCT 1380
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2095 TCATCTTTTGTGGGAGTCTGATATCCATGTGGATGCCAAGCAATTTGATTAACCTGGAG 2154
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1381 TCTTCTTTCTGTGTGGAGCTGATATCCAGTGGATGCCAAGCAATTTGATGTGGAG 1440
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2155 AATGCTTATGTTTTTATTTGACAGAGGTCAATTTTGGGGGGTGCACCGCTGTCTGGGTAT 2214
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1441 AACGCTTATGTTTACAGTGCACAGAGGCAATTTAGGGGGCACACCGCTGCCCTCATG---T 1496
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2215 TCATTTTTCATCAGCTGAGCATTTGTGATGCTATGCTTGTGGCTTCTCAGTTCATTAAGAG 2274
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1497 TTAGTTTTTACCATCGAGCACTGCTG---TTGCCCTCTTAGGTTTTTCACTCCATGAAG 1552
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2275 CAATAATGAAGTATTTAACTCTTTTCACTACAGTCTTTCGAAAGTATGCTATTTAAAT--- 2331
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1553 CAATAATGAAGTATTTTACTATTTTACAGGTCCCGCAAGAAATATCAAGTATGCTATT 1612
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2332 ----ACTTGGCCAGGTATTAATTTGCCAGTCAGTCTCTTTTATAGTGAGAAATTTATTTGGTT 2387
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1613 TGTCACTTGTCCAGGTATTAATTTACCAATCAGTCTCTGTACAAATAAGAAATATATCTG--- 1669
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2388 AGTAATATAAATATTTTAACTAAATATATAAATCTATAAATCTATAAATGTTTCAATTA 2447
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1670 -----TTAAACAATTTCAAGTGAGAGATCTAAGCCCTATAATGGTGA----ATCTTCATTA 1718
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2448 AAAGCATAGCACTTTGAAATTAACCTATATAAATAGCTCATATTTTACATTCAGCTTTTC 2507
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
|||||
1719 AAAG-----TTGGAATTAACGATATAAATATCTCATACTTAAAGTTGAGCGCTTTC 1770
|||||
2508 ATTGTATCAGTCTGAATCTTTTACGACTTTAAGAAATGACTATGATTAATATACCTG 2567
|||||
1771 ATTGTATTAAGTCTTAATCTTCACTGCTTGAGAAAAATGA--TTGCATTAATATACCTG 1828
|||||
2568 ACCATGAAGAAAAATGATACCTCAATGCAATGCAATTTGCACTGGTGATTCCACTGCACA 2627
|||||
1829 ACCATGGAAGAACTAAGTACCTC--AATGCAATGCAATTTGCACTGGTGCTGAGTGCACA 1887
|||||
2628 AATCTTTGTCATCTTGTATATAGGTATTTTTCATATGGGTTCACATGCGACACAAACCC 2687
|||||
1888 GGTCTGTCT-CATCATGTATAGGTATCTTT--ACATGTTGATAGGAACAC-GCACC 1943
|||||
2688 ATTTTCATTCAGTAGTAACCTTGGAGGCTGCGCATTTTCCATTTTACCAATCAACCAACCGCT 2747
|||||
1944 ATTTTCCTTTAGTATAAACC--TCAGATTGCTGCCATTTCTCACCTGACCCAAAGCCAGCT 2802
|||||
2748 GAAGTGAACTCGAACTGTTTTCATTAATCTTTCAAAAGTTGTTTACATCAATCTTA 2807
|||||
2003 GCAGATGAACCTCAAACTTGTCTCATAAACCATTTCAA-----GATATTA 2048
|||||
2808 AAATTTCAAATGCTGCAGGCTAATTTAATGATATAAAATATTAGTAAGAAAAAGTATGTA 2867
|||||
2049 AAGTTCCAGAAATGCTGCAGGCTAATTTAATGATATAAAATATTAGTATTTGTAAG--AAGTATAT 2105
|||||
2868 TTGCATCTTATAGTAGATAGATCAACATACAAATTCATTCAGTGCATGCTTTAGGTG 2927
|||||
2106 TCGCATATATAGTAGTAGATCAGAAGTATCAATTTGACTCAATGATGCTTTAGGT 2165
|||||
2928 TTAACATGAGATTGTAAGTTTACTGTTAGTCTTGGCATCTGCTGCTAGTGAGT 2987
|||||
2166 TTAAGCATGAGATTGTACATGTTTCACTGTTAAGTCTTTGCATC--TGTGCTAGGTGAG- 2222
|||||
2988 ATGAGAGATGTCRAGGACTGACGCTATTTGTCCTTAAAAAAAGGCTGTTTGTAG 3047
|||||
2223 ----GGAGATGTTAAGGACATAAAATATTTGTCCTTAAAAAAAGGCTGTTTGTAG 2277
|||||
3048 GGGTTTAAATATGCTTATTTTGTGTGCTCTCTCACTACTATTTACACATGTTGCTTTGT 3107
|||||
2278 ACATTTGATGATGCTTATTTAGTGTCTCACTACCTAGTAAACATAAGTGTCTACTACC 2337
|||||
3108 GGGTTTGTGTTGTATGTCGCTGTTTATACAGTAGTTAAATTTCCATGCAAGAAAAATAA 3167
|||||
2338 TAGTAAACAT--AATTGCTGTGTTGTCATAGTAGTTAAATTTCCATGCAAGAAAAATAA 2395
|||||
3168 TGTCTCTGAATTTCTCAA 3184
|||||
2396 TGTCTCTGAATTTCTCAA 2412
|||||

RESULT 6
AK051728
LOCUS
DEFINITION
Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
enriched library, clone:D130067011 product:hypothetical
Glycerophosphoryl diester phosphodiesterase/Glycosyl hydrolase,
starch-binding domain containing protein, full insert sequence.
ACCESSION
AK051728
VERSION
AK051728.1 GI:26342153
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
Carninci, P. and Hayashizaki, Y.
AUTHORS
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
9279253
MEDLINE
10349636
PUBMED
REFERENCE
2
```

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

20493374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 2228)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, I., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Ozaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. .2228
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:D130067011"
/db_xref="taxon:i10090"
/clone="D130067011"
/tissue_type="spinal ganglion"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
103. .1293
/note="unnamed protein product; hypothetical

CDS

Glycerophosphoryl diester phosphodiesterase/Glycosyl hydrolase, starch-binding domain containing protein (InterPro IPR004129, InterPro IPR002044, evidence: putative)
/codon_start=1
/protein_id=BAC34739.1
/db_xref=GI:26342154
/translation="MSRSRKRTIGKVRVDFIIKPLPGVSCMSQSFYSKWKPRIPLD VHRGAGNSTTTAKLAKVQNTIASHRNAASHGAFAVEFDVHLSKDFVYVYHDIJTC LTKWKYADPEIPVKEITFDQQLKLSHTALTKDKRKOLYEEENFFSENO PPSLKLWLESPELVGNIELKWLCOHRDGVNDGNLSTYEDMAYFLDILILKVLSEN GKRIYVSSFDADI CTMRQKQNPILFLTOKSDIYFELMDLSRTPIAMSAQF ENILGINATEDLNPSTVQBAKAGLWIFCWGDNDPNNRKLKKEFGVNGLIYDR IYDMPEQNPISFVQELERKQELPELKNCLPFTVSHFIPSSFCVEPDHVDANGIDS VENA"

ORIGIN

Query Match 44.3%; Score 1420.2; DB 3; Length 2228;
Best Local Similarity 81.9%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 343; Indels 72; Gaps 18;
QY 871 GTTCAGGGTGATGCCCTTCCTGACATGCGGTACAGCTGTCTCTTATCATCCACCAT 930
DB 1 GTTCAGGGTGATGTTCTTCCTGACACGCTGGGCACAGCATGCCCTCTGCTCTACCAT 60
QY 931 GCTGAGGTGGAAGAGTCTCGAATCTTACTCTTCCCATCATGACAGAGAAATTCCTGG 990
DB 61 GCTGAGGTGGAAGAGCGCTGGAATCTTACTCTTCCCATCATGACAGAGAAATTCCTGA 120
QY 991 AAAACAATAGGCAAGTGAGTTGATATATAATTAATTAAGCAATTAACAGATACAGT 1050
DB 121 AAAACATATAGGCAAGTCAGATTGATTTATCATCATCAAGCAATTAACCTGGATATAGT 180
QY 1051 TGTGACATCAATCTTCAATTTTCCAGTATTCGAGCCAGAAATCAATCAATGTTGGC 1110
DB 181 TGTCTATGACGTCCTTCAATTTTCCAGTATTCGAGCCAGAAATCAATCAATGTTGGC 240
QY 1111 CATGAGGTGCGAGGAACCTCAACAACACTGCCAGCTGGCTGAAAGTTCAAGAAATATCT 1170
DB 241 CATGCTGTGCGAGGAACCTCAACAACACTGCCAGCTGAAAGTTCAAGAAATATCT 300
QY 1171 ATTGCTCTTTAAGAAATGCTGTAGTCAATGCTGAGCTGCTGAGAAATTTGACGTACAC 1230
DB 301 ATCGCTCTTTAAGAAATGCTGTAGTCAATGCTGAGCTGCTGAGAAATTTGACGTACAC 360
QY 1231 CTTTCAAGAGCTTTGTCCTGGTGTATATCATGATCTTACCTGTTGTTGACTATGAA 1290
DB 361 CTTTCAAGAGCTTTGTCCTGGTGTATATCATGATCTTACCTGCTGCTGACCAATGAG 420
QY 1291 AAGAAATTTGATGCTGATCCAGTTGAATTTATTTGAATTTCAAGTAAAGAAATTAACATTT 1350
DB 421 AGAAATATGAAGCTGATCCAGTTGAATTTGTTGAATTTCCAGTAAAGAAATTAACATTT 480
QY 1351 GACCAACTCCAGTTGTTAAGCTCACTCATGCTGACCTGCACTGAAATCTAAGGATCGGAA 1410
DB 481 GACCAACTCCAGTTGTTAAGCTCTCTCATGCTGACCTTAAACCAACCAAGACCGGAA 540
QY 1411 GAATCTGTGGTTCAGGAGGAAATTCCTTTTCAGAAATTCAGCAATTTCTCTCTTAAG 1470
DB 541 CAATCTTTGTATGAGGAGGAAATTTCTTTTCTGAAATTCAGCAATTTCTCTCTTAAG 600
QY 1471 ATGGTTTTAGAGTCTTTGCGAGAGATGAGGTTTAAATTTGAATTTGAATAAATGGATCTGC 1530
DB 601 ATGGTTTTAGAAATTCAGGAGGAAATGAGGATTTAATATAGAAATAAATGGATTTGC 660
QY 1531 CAGCAAGGAGATGGAATGTTGGATGTAATCTTATCAACATATTTTGAATGAATCTGTTT 1590
DB 661 CAACACAGGATGAGATGTTGGATGCAATCTTATCAACATATTTTGAATGAATGTTT 720
QY 1591 TTGGATATTAATTTTAAACCTGTTTGAATAATTCGGGAGAGGAGGATGATGTTTCT 1650
DB 721 TTGGATATTAATTTTAAACCTGTTTGAATAATTCGGGAGAGGAGGATGATGTTTCT 780

QY 1651 TCATTGTATGCAGATATTTTGCACAATGGTTGGCAAAAGCAGAAACAAATATCCGATACTA 1710
DB 781 TCTTTTGTATGCAGATATTTGTACAATGGTTGGCAAAAGCAGAAACAAATATCCGATACTA 840
QY 1711 TTTTAACTCAGGAAATCTGAGATTTATCTGAACTCATGAGCTCAGATCTCGGACA 1770
DB 841 TTTTGAACCAAGGAAAGTCTGATATTTTACCCGAACTCATGAGCTCAGATCTCGGACA 900
QY 1771 ACCCCATTTGCAATGAGCTTTTGCAAGTTTGAATTTACTTGGGATAAATGTACATCT 1830
DB 901 ACACCCATTTGCAATGAGTTTGCACAGTTTGAATATTTTGGGATAAATGCCCATCT 960
QY 1831 GAAGACTTCTCAGAAAACCCATCTTATATCAAGAGGCAAAAGCTAAGGAGACTAGTCATA 1890
DB 961 GAAGACCTCTTAGAAAACCCATCTTATGTCCAAGAGGCAAAAGCTAAGGAGTTGGTCATA 1020
QY 1891 TTCTGCTGGGTGATGATACCAATGATCTGAAACAGAAAGAAATTCAGGAACTTGGGA 1950
DB 1021 TTCTGCTGGGTGATGATACCAACGATCTGAAACAGAAAGAAATTCAGGAAATTTGGA 1080
QY 1951 GTTAATGGTCTAAATTTATGATAGGATATATGATTCGATGCTGAAACCAAAATATATC 2010
DB 1081 GTAAATGGTCTAAATTTATGATAGGATATATGATTCGATGCTGAAACCAAAATATATC 1140
QY 2011 CAAAGTGGAGCAATTTGGAACGCTGAAAGCAGGAAATTCAGAGCTTTAAGAGCTGTTGT 2070
DB 1141 CAAAGTGGAGCAATTTGGAACGCTGAAAGCAGGAAATTCAGAGCTTTAAGAGCTGTTGT 1200
QY 2071 CCCACTGTAGCGCTTTGTTCCCTCATCTTTGCTGGGAGTCTGATATCCATGTTGAT 2130
DB 1201 CCCACTGTAGCGCACTTCACTCCCTCTTCTTCTGTTGGAGCTGATATCCACGTTGAT 1260
QY 2131 GCCAAGCAGCAATTTGAACTGAGAGATCTTAGTTTATTTGACAGAGGTCATTTTGGG 2190
DB 1261 GCCAAGCAGCAATTTGAGTGTGGAGAACCTTAGTTTCAAGTGCACAGAGGACATTTAGG 1320
QY 2191 GCGGTGACCGCTGTTTGGGTATTTTCACTCATCTGAGCATTTGTTGATCTATGCT 2250
DB 1321 GCACACCGCTGCGCTCATG---TTAGTTTTCACCATCGAGCACTGCTG---TTGCT 1372
QY 2251 TTTGGCTTCTCAGTTCAATGAGCAATTAATGAGTATTTAATCTTTTCACTACAGTTCT 2310
DB 1373 CTTAGTTTCTCAGTCCCAATGAAGCAATTAATGAAGTATTTTACTATTTTCACTACAGTTCT 1432
QY 2311 TGCAAGTATGCTATTTAAAT---ACTTGGCCAGGTATAATTTGCCAGTCAGTCTCT 2363
DB 1433 CGCAAGAAATATCAAGTATGCTATTTGTCACCTTGTCCAGGTATTAATACCAATCAGTCTCT 1492
QY 2364 TTATAGTGAAGAAATTTATTTGGTTAGTAAATATAATTTTAAACTATAATATAATCT 2423
DB 1493 GTACAATAAGAAATATACTG-----TTAAACATTTTCAGTGAGATATCTAAGCT 1542
QY 2424 ATAATGTTAAACATATGTTTCAATTAAGCATAGCACTTTGAAATTAATTAATAATAGC 2483
DB 1543 ATAATGTTGCA---ATCTTCAATTAAG---TTGGAATTAAGCATATAATATC 1590
QY 2484 TCATATTTTACACTTACAGCTTTTCAATTTGATCAGGTCTGAAATCTTTAGCACTTAAGAA 2543
DB 1591 TCATACTTAAAGTTTCAAGCTTTTCAATTTGATTAAGTCTTAATCTTCAGTGTCTGAGAA 1650
QY 2544 AATGACTATGCAATTAATATACCTGCAATGAAATAAATTAAGTACCTCAAAATGCATGCAAT 2603
DB 1651 AATGA---TTGCATATTAATACCTGCAATGAAATAAATTAAGTACCTC---AATGCATGCAAT 1707
QY 2604 TGCATGTTGATTCCAACTGCAAAATCTTTGTGGCCATCTTGTATATAGGATTTTATAC 2663
DB 1708 TGCATGTTGGCTCCAGCTGCAAGCTGTGTGT---CATCCATGTACATAGGTATCTTT---A 1764
QY 2664 ATGGGTTCACTGACACCAACCACTTTTCAATTCAGTATGACCTTTGAGGCTGCTGCAT 2723
DB 1765 CATGGTTGATAGGAACAG---CACCAATTTTCTTTTATAGTATAAACCT---TCAGATTGCTGCCAT 1822

Qy	2724	TTTTCCACTTAACCAACGAGCTGAAGGTGAACCTCGAAACTTGTGTTTCATAAAATCTTTC	2783
Db	1823	TTCTCACTGACCCAGCAGCCTGCAGATGAACCTCAAACTTGTCTCATPAACCAATTC	1882
Qy	2784	AAAAGTTGTTTTACATCAATGTTAAATTTCAAAATGCTGCAGGGTAATTTTAATGTATAA	2843
Db	1883	AAA-----GATATTAAAGTTCCAGAATGCTGCAGGGTAATTTAATGTATAA	1928
Qy	2844	AATATTAGTAGAAMAAAGTATTGTATTCATCTTAGTATAGATAGATACACATACAAAT	2903
Db	1929	AGTATTGTTAAG--AAGTATATATTGCATATATAGTAGTAGTGTAGATCAGAAGGTATCAAT	1985
Qy	2904	TCAATTCAAGTCAGTCCTTTAGTGTTAAGCATGAGATTGTACATGTTTACGTGTAGTGCC	2963
Db	1986	TTGACTCAATGATGCTTTAGGTTTTAAGCAAGATGTTACATGTTTCTTACGTTAAAGTCC	2045
Qy	2964	TTGCATCTGTGCTGCTAGTGTAGTATGAGAAGATGTCGAAGACTCGACGTATTTTGTTCG	3023
Db	2046	TTGCATC--TGGTCTAGTTCAG-----GGAGATGTTAAGGACATATAAATTTTGTTCG	2098
Qy	3024	CTAAAAAAAAGGCTGTTGTAGCGTTTTAAATATGCTTATTTTGTGTCTCTCACT	3083
Db	2099	CTAAAAAACAAGTCTGTTT-TAGACACTTTCAGTATGCTTATTTAGTGTCTCACTACCT	2157
Qy	3084	ACCTATTACACACTGTGCTTTGTGGTGTGTTTGTATGTGCTGTGTATACAGTAGT	3143
Db	2158	AGTAACATAGTGTCTCACTACCCAGTAAACAT--AATTGCTGTGTGTCATAGTAGT	2215
Qy	3144	TAAATTTCCATGC	3156
Db	2216	TAAATTTCCATGC	2228
RESULT 7			
AK049491			
LOCUS	AK049491	1618 bp	mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus 7 days embryo whole body cDNA, RIKEN full-length		
	enriched library, clone:C430017C03 product:hypothetical		
	Glycerophosphoryl diester phosphodiesterase/Glycosyl hydrolase,		
	starch-binding domain containing protein, full insert sequence.		
ACCESSION	AK049491		
VERSION	AK049491.1	GI:26340223	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1		
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,		
	Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,		
	Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,		
	Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,		
	Yamamoto,K., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwaki,K.,		
	Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,		
	Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,		
	Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format		
	sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		

PUBMED	11076861		
REFERENCE	4		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the		
	FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE	5		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research		
	Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation		
JOURNAL	of 60,770 full-length cDNAs		
REFERENCE	Nature 420, 563-573 (2002)		
AUTHORS	6 (bases 1 to 1618)		
	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,		
	Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,		
	Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,		
	Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,		
	Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,		
	Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,		
	Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,		
	Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,		
	Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,		
	Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaira,S.,		
	Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,		
	Muramatsu,M. and Hayashizaki,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of		
	Physical and Chemical Research (RIKEN), Laboratory for Genome		
	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),		
	RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,		
	Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,		
	URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,		
	Fax:81-45-503-9216)		
COMMENT	cDNA library was prepared and sequenced in Mouse Genome		
	Encyclopedia Project of Genome Exploration Research Group in Riken		
	Genomic Sciences Center and Genome Science Laboratory in RIKEN		
	Division of Experimental Animal Research in Riken contributed to		
	prepare mouse tissues.		
	Please visit our web site for further details.		
	URL:http://genome.gsc.riken.jp/		
	URL:http://fantom.gsc.riken.jp/		
FEATURES	Location/Qualifiers		
source	1. .1618		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="FANTOM DB:C430017C03"		
	/db_xref="taxon:10090"		
	/clone="C430017C03"		
	/tissue_type="whole body"		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		
	/dev_stage="7 days embryo"		
	129, _1319		
CDS	/note="unnamed protein product; hypothetical		
	glycerophosphoryl diester phosphodiesterase/Glycosyl		
	hydrolase, starch-binding domain containing protein		
	(InterPro IPR004129, InterPro IPR002044, evidence:		
	InterPro)		
	putative"		
	/codon_start=1		
	/protein_id="BAC33775.1"		
	/db_xref="GI:26340224"		
	/translations="MSRNSRKTIGKVRVDFIILKPGVSCSMQSSFSKYWKPRIPLD		
	VGHRGAGNSTTTAKLVQENTIASLRNAAASHGAAVFEDVHLKQFVPPVYHDLTCC		
	LTKRKYEADPVELFELPVKELTFDQOLKLSHTALTKDKRKSLYEENPSENO		
	PFPLKWLSELPENFGNIQIKWICQHRDVGWGNLSTYFDNNVLDILKTVSQNS		
	ENRIFVFSFDADICTVVRQKNQKPIFLITQKSDIYPELMDLRKRTPIAMSFQF		
	GNRIINATDEDLIRNPVYQEAAGLVIFCGDDTNDPNERKLKEFGVGLIYDR		
	IYDWMPEQPNIFQVEQLERLKLQELPELKNCLCTVSHFIPSSFCVPEPDIVHDANGIDS		
	VENA"		
ORIGIN			
Query Match	36.5%;	Score 1170.8;	DB 3; Length 1618;

Best Local Similarity 84.1%; Pred. No. 7.6e-261; Matches 1361; Conservative 0; Mismatches 242; Indels 15; Gaps 3;	
QY	845 TCGAAGAAGATCTCAGTGAACAGTGTAGTTCAGGGTGTATGCCCTTCTCTGACATGTGGGTA 904
Db	1 TTGAGGAAGATCTCAGTGAACAGTGTAGTTCAGGGTGTATGCCCTTCTCTGACATGTGGGCA 60
QY	905 CAGCTGTCTCTTATCATCATCAGTGTGTGAGAGTGAAGAGTGTGGAATCTTACTC 964
Db	61 CAGCTGTCTCTTATCATCATCAGTGTGTGAGAGTGAAGAGTGTGGAATCTTACTC 120
QY	965 TTCCCATCATGACAGCAAAATTCGCCGAACAAATAGSCAAAGTGAGAGTGTGACTATATA 1024
Db	121 TTCCCATCATGACAGCAAAATTCAGAAAACAAATAGSCAAAGTGAGAGTGTGACTATATA 180
QY	1025 TTATTAAGCCATTAACAGGATACAGTGTGTGACATGAATCTTCAATTTTCCAAAGTATTTGA 1084
Db	181 TCATCAAGCCATTAACCTGGATATAGTGTGTGTATGTCAGTCTTCATTTTCCAAAGTATTTGA 240
QY	1085 AGCCAGAATACCATTTGGATGTGGGATCGAGTGCAGGAACCTCTACAACTGCCC 1144
Db	241 AACCAAGAATACCATTTGGAGTGTGGACATCGTGTGTCAGGAACTCAACAACTGCCC 300
QY	1145 AGCTGCTAAAGTTCAGAAATACTATTGCTTCTTTAAAGAACTGCTAGTCAATGTTG 1204
Db	301 AGCTGCTAAAGTTCAGAAATACTATTGCTTCTTTAAAGAACTGCTAGTCAATGTTG 360
QY	1205 CAGCTTTGTAGAAATTTGACGTACACCTTTCAAGGACCTTTGTCGCCGTGTATATCATG 1264
Db	361 CAGCTTTGTAGAAATTTGATGTCACCTTTCAAGGACCTTTGTCGCCGTGTATATCATG 420
QY	1265 ATCTTACCTGTTGTTGACTATGAAGAAATTTGATGCTGATCAGTTGAATTTTGG 1324
Db	421 ACCTCACCTGCTGCTGACCATGAAGAGGAAATATGAAGCTGATCCAGTTGAATTTGTTG 480
QY	1325 AAATCCAGTAAAGAAATTAACATTTGACCACTCCAGTGTGTAAGCTCACTCATGTA 1384
Db	481 AAATCCAGTAAAGAAATTAACATTTGACCACTCCAGTGTGTAAGCTCACTCATGTA 540
QY	1385 CTGCACTGAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTTTCAG 1444
Db	541 CTGCACTGAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTTTCAG 600
QY	1445 AAATCAGCAATTTCTTCTTTAAGATGTTTGTAGAGTCTTTGTCAGAGATGTAGGT 1504
Db	601 AAATCAGCAATTTCTTCTTTAAGATGTTTGTAGAGTCTTTGTCAGAGATGTAGGT 660
QY	1505 TTAACATTTGAATTAATGATCTGCAGCAAGGATGGAATGCTGGATGTAATTTAT 1564
Db	661 TTAATATACAAATTAATGATTTGCCACACAGGATGGATGGATGGATGGCACTTAT 720
QY	1565 CAACATATTTGACATGAATCTGTTTTCGATATAATTTTAAAGCTGTTTTCAGAAATTT 1624
Db	721 CAACATATTTGATATGAATGTTTTCGATATAATTTTAAAGCTGTTTTCAGAAATTT 780
QY	1625 CTGGGAAGAGGAGATAGTGTCTTCTTCAATTTGATCAGATTTGCAAAATGTTTCGGC 1684
Db	781 CTGGGAACAGGAGATAGTGTCTTCTTCAATTTGATCAGATTTGCAAAATGTTTCGGC 840
QY	1685 AAAGCAGAACAAATATCCGATCTATTTTAACTCAAGGAAATCTGAGATTTATCTTG 1744
Db	841 AGAAGCAGAACAAATATCCGATCTATTTTAACTCAAGGAAATCTGAGATTTATCTTG 900
QY	1745 AACTCATGACCTCAGATCTCGGACAAACCCCAATTCATGAGTGTTCACAGTTTGA 1804
Db	901 AACTCATGACCTCAGATCTCGGACAAACCCCAATTCATGAGTGTTCACAGTTTGA 960
QY	1805 ATCTACTGGGGATAAATGTATCACTATGAAGCTTGTCTGAGAACCCATCTATATTTCAAG 1864
Db	961 ATATTTGGGGATAAATGTATCACTATGAAGCTTGTCTGAGAACCCATCTATATTTCAAG 1020
QY	1865 AGGCAAAAGCTAAGGAGTGTATGATCTGCTGGGTGATGATCAATGATCTCTGAA 1924
Db	1021 AGGCAAAAGCTAAGGAGTGTATGATCTGCTGGGTGATGATCAATGATCTCTGAA 1080
QY	1925 ACAGAAGGAAATTTGAAGAACTTTGAGTGTAAATGGTCTAAATTTATGATAGATATGAT 1984
Db	1081 ACAGAAGGAAATTTGAAGAACTTTGAGTGTAAATGGTCTAAATTTATGATAGATATGAT 1140
QY	1985 GGATCGCTGAAACAAACAAATATATTTCCAAAGTGGAGCAATTTGGAACGCTGGAAGCAAT 2044
Db	1141 GGATCGCTGAAACAAACAAATATATTTCCAAAGTGGAGCAATTTGGAACGCTGGAAGCAAT 1200
QY	2045 TCCAGAGCTTAAAGAGCTGTTTGTGTCACCTGTTAGCGCTTGTTCCTCATCTTTCT 2104
Db	1201 TCCAGAGCTTAAAGAGCTGTTTGTGTCACCTGTTAGCGCTTGTTCCTCATCTTTCT 1260
QY	2105 GTGGGAGTCTGATATCCATGTGATGCAACCGCAATTTGATACCTGGAGAAATGCTTGT 2164
Db	1261 GTGGGAGTCTGATATCCATGTGATGCAACCGCAATTTGATAGTGTGGAGAAACGCTTGT 1320
QY	2165 TTTTATTTGACAGAGTCAATTTTGGGGGGTGCACCGCTGTTCTGGGTATTTCAATTTTCA 2224
Db	1321 TCACAGTGCACAGAGGACATTTTAGGGGGCACACCGCTGCCCTCATG----TTTAGTTTTC 1376
QY	2225 TCACAGTGCATGTTGATCTATGCTTCTTGGGCTTCTCAGTTCATGAGCAATTAATGAA 2284
Db	1377 CCATCGAGCACTGCTG----TTGCCCTCTTAGGTTTTCAGTCCCAATGAGCAATTAATGAA 1432
QY	2285 GTATTTAACTCTTTTCACTACAGTCTTCTGCAAGTATGCTA-----TTTAAATTAATCTGG 2337
Db	1433 GTATTTAACTCTTTTCACTACAGTCTTCTGCAAGTATGCTA-----TTTAAATTAATCTGG 1492
QY	2338 CCAGGTAAATTTGACAGTCACTCTTTTATAGTGAAGAAATTTATGTTAGTATATAA 2397
Db	1493 CCAGGTAAATTTTACCAATCAGTCTCTGTACAATAAGAAATATATCTGTTAAACATTTTCAG 1552
QY	2398 ATATTTTAACTAAATATATAATCTATAATGTTTAAACATATGTTTCATTAAGGATA 2455
Db	1553 TGAGATATCTAAGCCTATAATAGTGGAATCTTCAATTAAGGATTAACGATATA 1610
RESULT 8	
BX344122/c	
LOCUS	BX344122 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION	clone CS01059YA01 3-PRIME, mRNA sequence.
ACCESSION	BX344122
VERSION	BX344122.2 GI:46262755
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	On May 1, 2003 this sequence version replaced gi:30309075. Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of invitrogen. This sequence belongs to sequence cluster 9838.r For more information about this cluster, see http://www.genoscope.cns.fr/cdna?e=CS1A1015ZE0INP1&c=9838.r.
FEATURES	Location/Qualifiers
source	1..998 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606"

[illegible]

Qy	3048	GC	GGTTTAAATAGCTTATTTCTGCTGCTCTCACTACCTATTACACACGTTGCTTTGT	3100
Db	98	GC	GGTTTAAATAGCTTATTTCTGCTGCTCTCACTACCTATTACACACGTTGCTTTGT	39
Qy	3108	GG	GGTTTCTTTGTATGTGCGTGTATTATACAGTAGTTA	3145
Db	38	GG	GGTTTCTTTGTATGTGCGTGTATTATACAGTAGTTA	1
RESULT 9				
LOCUS	BX344123		1006 bp	mrna linear EST 07-APR-2004
DEFINITION	BX344123 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI059YA01 5-PRIME, mRNA sequence.			
ACCESSION	BX344123			
VERSION	BX344123.2		GI:46264695	
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 1006)			
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
JOURNAL	Full-length cDNA libraries and normalization			
COMMENT	Unpublished (2001)			
	On May 1, 2003 this sequence version replaced gi:30311247.			
	Contact: Genoscope			
	Genoscope - Centre National de Sequencage			
	2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE			
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9838.f			
FEATURES	For more information about this cluster, see			
source	http://www.genoscope.cns.fr/cdna?s=CSIAI0152E01QP1&c=9838.r.			
	Location/Qualifiers			
	1..1006			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="CSODI059YA01"			
	/tissue_type="PLACENTA COT 25-NORMALIZED"			
	/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"			
	/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN				
Query Match	28.7%	Score	921.4;	DB 5; Length 1006;
Best Local Similarity	97.0%;	Pred. No.	6.4e-203;	
Matches	950; Conservative	8; Mismatches	19; Indels	2; Gaps 2;
Qy	1144	CAGCTGCTGAAGTTCAGAAATACATTATGCTTCTTTAAGAATGCTGCTAGTCATGGT	1203	
Db	1	CAGCTGGTACCGTCCGGAAATCCCGGATCTTCTTTAAGAAATGCTGCTAGTCATGGT	60	
Qy	1204	GCAGCCCTTTGTAGAAATTTGACGTACACCTTTCAAAGGACCTTTGTGCCCGGTGATATCAT	1263	
Db	61	GCAGCCCTTTGTAGAAATTTGACGTACACCTTTCAAAGGACCTTTGTGCCCGGTGATATCAT	120	
Qy	1264	GATCTTACCTGTTGTTTGCATATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTTT	1323	
Db	121	GATCTTACCTGTTGTTTGCATATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTTT	180	
Qy	1324	GAAATCCAGTAAAGAAATTAACATTTTGACCACTCCAGTTGTTTAAAGCTCACTCATGTG	1383	
Db	181	GAAATCCAGTAAAGAAATTAACATTTTGACCACTCCAGTTGTTTAAAGCTCACTCATGTG	240	
Qy	1384	ACTGCACCTGAAATCTAAGGATCGGAAAGAAATCTGTGTTCCAGGAGGAAATTCCTTTTCA	1443	

QY 2128 GATGCCAACGCGATGATTAACGTGGAGATGCTTAGTCTTTTATTTGACAGAGGTCAATTTT 2187
|||||
Db 721 GATGCCAACGCGATGATTAACGTGGAGATGCTTAGTCTTTTATTTGACAGAGGTCAATTTT 780
|||||
QY 2188 GGGGGCGTGCAACCGCTGCTTGGGGTATTCATTTTTCATCTCACTGAGCATTTGTCATCTATG 2247
|||||
Db 781 GGGGGCGTGCAACCGCTGCTTGGGGTATTCATTTTTCATCTGAGCATTTGTCATCTATG 840
|||||
QY 2248 CC-TTTTGGGCTTCTCAG-TTCAATGAAGCAATAA 2280
|||||
Db 841 CTTTTTGGGCTTCTCAGTTTCAATGAAGCCATAA 875
|||||

RESULT 11
BM478430 1128 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6508010 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5578502
DEFINITION 5', mRNA sequence.
ACCESSION BM478430
VERSION BM478430.1 GI:18527472
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1128)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12333 row: k column: 15
High quality sequence stop: 618.
Location/Qualifiers
1. 1128
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5578502"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

FEATURES
source
Query Match 26.0%; Score 835; DB 4; Length 1128;
Best Local Similarity 91.6%; Pred. No. 7.8e-183;
Matches 1008; Conservative 0; Mismatches 71; Indels 21; Gaps 11;
QY 803 AGACGATGGAACCCAGATAACCTGGAACTAATCTTTGATTTTTCGAGAAATCTCAGTG 862
|||||
Db 13 AGACGATGGAACCCAGATAACCTGGAACTAATCTTTGATTTTTCGAGAAATCTCAGTG 72
|||||
QY 863 AGACGATGTTAGGGTATGCCCTTCTGGACATGTTGGTACAGTTCCTCTTATCAT 922
|||||
Db 73 AGCAGCTAGTTTACGGGTGATGCCCTTCTGGACATGTTGGTACAGTTCCTCTTATCAT 132
|||||
QY 923 CCACCATTTGCTGAGAGTGGAAAGAGTGGTAAATTTCTTACTTTCCCATCATGAGCAGAA 982
|||||
Db 133 CCACCATTTGCTGAGAGTGGAAAGAGTGGTAAATTTCTTACTTCTCCCATCATGAGCAGAA 192
|||||
QY 983 ATTCCCGGAAACAATAGGCAAGTGGAGTTGACTATATAATTATTAAAGCCATTACCAG 1042
|||||

Db 193 ATTCGCGGAAAAAATAGCAAAAGTGAGAGTTGACTATATAATTATTAAAGCCATTACCAG 252
|||||
QY 1043 GATACAGTTGTGACATGAATCTTTCATTTTCCAGATTTTGAAGCAAGAATACCATTTGG 1102
|||||
Db 253 GATACAGTTGTGACATGAATCTTTCATTTTCCAGATTTTGAAGCAAGAATACCATTTGG 312
|||||
QY 1103 ATGTTGGCCATCGAGTGCAGGAAACTCTTACAACTGCCAGCTGGCTAAAGTTCAAG 1162
|||||
Db 313 ATGTTGGCCATCGAGTGCAGGAAACTCTTACAACTGCCAGCTGGCTAAAGTTCAAG 372
|||||
QY 1163 AAAATACTATTGCTTCTTTTAAAGAAATGCTGTAGTCAATGCTGCAGCTTTGTAGAAATTG 1222
|||||
Db 373 AAAATACTATTGCTTCTTTTAAAGAAATGCTGTAGTCAATGCTGCAGCTTTGTAGAAATTG 432
|||||
QY 1223 AGCTACACCTTTCAAAGGACTTTTGGCCGTGGTATATCATGATCTTACCTGTGTTGA 1282
|||||
Db 433 AGCTACACCTTTCAAAGGACTTTTGGCCGTGGTATATCATGATCTTACCTGTGTTGA 492
|||||
QY 1283 CTATGAAAAAGAAATTTGATGCTGATCCAGTTTGAATTTTGGAAATTTCCAGTAAAGAAAT 1342
|||||
Db 493 CTATGAAAAAGAAATTTGATGCTGATCCAGTTTGAATTTTGGAAATTTCCAGTAAAGAAAT 552
|||||
QY 1343 TAACTTTGACCAACTCCAGTTTGTAAAGCTCACTCATGCTGACCTGCAATCTAAGG 1402
|||||
Db 553 TAACTTTGACCAACTCCAGTTTGTAAAGCTCACTCATGCTGACCTGCAATCTAAGG 612
|||||
QY 1403 ATCGGAAAGAAATCTGCTGTTTCCAGGAGGAAATTTCCCTTTTTCAGAAATTCAGCATTTCCTT 1462
|||||
Db 613 ATCGGAAAGAAATCTGCTGTTTCCAGGAGGAAATTTCCCTTTTTCAGAAATTCAGCATTTCCTT 672
|||||
QY 1463 CTCTTAAGATGGTTTGTAGAGTC-TTTGCCAGAAAGATGTAGGG-TTTAAACATTTGAAATAAA 1520
|||||
Db 673 CTCTTAAGATGGTTTGTAGAGTC-TTTGCCAGAAAGATGTAGGG-TTTAAACATTTGAAATAAA 732
|||||
QY 1521 ATGGATCTGCAGCAAAAGGATGGAAT-GTGGGATGCTAACTTATCAACATATTTTGACA 1579
|||||
Db 733 ATGGAICTGCCAGAAAGGATGGAATGGGGGATGTTAACTTATTTCACATATTTTGACA 792
|||||
QY 1580 TGAATCTGTTTTTGGATATAATTTTAAAACTGTTTTAGAAATTTCTGG---GAAGAGGA 1636
|||||
Db 793 TGAATCTGTTTTTGGATATAATTTTAAAACTGTTTTTAAAAATTTCTGGGGAGAGAGGA 852
|||||
QY 1637 GAATAGTGTCTTCTCATTTTGTGATGAGATATTTTGCAAA--TGGTTCCGGCAAAAGCAGAA 1694
|||||
Db 853 AAAAAGTTTTTCTCATTTTGTATGAGATTAATTTGCCACTGGGTTCGCAAAAGCCGAAA 912
|||||
QY 1695 CAAATATCC---GATACTATTTTAACTCAAGGAAATCTG-AGATTATCTCTGAATCTC- 1749
|||||
Db 913 CAATTATCCCATCTACTAATTTTCTTCTCAAGGAAATCTGAAATTTATCCCGGACTCC 972
|||||
QY 1750 -ATGGACCTCAGATCTCGGCAACCCCATTTGCAATGAGCTTTTGCACAGTTTGAATCT 1808
|||||
Db 973 ATGGAATCTCAGAACTCGGACAAACCCCATTTGCAATGAGCTTTTGGCCGCTTTGAAAAAT 1932
|||||
QY 1809 AC----TGGGGATAAATGTACATAC--TGAAGACTTGTCTCAGAAA-CCCATCTTATATTC 1861
|||||
Db 1033 TCCACTGGGGATAAAGGTACCTACCTGGAGAACTTGTCTCAAAACCCCTTCCCAATTTTC 1092
|||||
QY 1862 AAGAGGCAAAAGCTAAAGGGA 1881
|||||
Db 1093 AAAAGGCAAAAGTCTAAGGA 1112
|||||

RESULT 12
B0049943
LOCUS AGENCOURT_7048600 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5787913
DEFINITION 5', mRNA sequence.
ACCESSION B0049943
VERSION B0049943.1 GI:19809283
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 973)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12879 row: a column: 02
High quality sequence stop: 735.
Location/Qualifiers
1. 973
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5787913"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

Query Match 25.5%; Score 816; DB 5; Length 973;
Best Local Similarity 99.9%; Pred. No. 2e-178; Indels 1; Gaps 1;
Matches 827; Conservative 0; Mismatches 0

QY 1808 TACTGGGATAATGTACATCTGAGACTTGCTCAGAACCCATCTATATTCAGAGG 1867
DB 1 TACTGGGATAATGTACATCTGAGACTTGCTCAGAACCCATCTATATTCAGAGG 60

QY 1868 CAAAAGCTTAAGGACTAGTCATATTCCTGCTGGGTGATGATCAATGATCCTGAAACA 1927
DB 61 CAAAAGCTTAAGGACTAGTCATATTCCTGCTGGGTGATGATCAATGATCCTGAAACA 120

QY 1928 GAAGGAATTAAGGACTAGTCATATTCCTGCTGGGTGATGATCAATGATCCTGAAACA 1987
DB 121 GAAGGAATTAAGGACTAGTCATATTCCTGCTGGGTGATGATCAATGATCCTGAAACA 180

QY 1988 TGCCTGAACACCAATATATTCAGTGGAGCAATTTGGAACGCTGAAGCAGGAATTC 2047
DB 181 TGCCTGAACACCAATATATTCAGTGGAGCAATTTGGAACGCTGAAGCAGGAATTC 240

QY 2048 CAGAGCTTAAGAGCTGTTTGTGTCACATGTTAGCGCTTGTTCCTCATCTTTGTGTG 2107
DB 241 CAGAGCTTAAGAGCTGTTTGTGTCACATGTTAGCGCTTGTTCCTCATCTTTGTGTG 300

QY 2108 GGGAGTCTGATATCCATGTTGGATGCCACCGCATGATCAACGTGGAGATGCTTAGTTT 2167
DB 301 GGGAGTCTGATATCCATGTTGGATGCCACCGCATGATCAACGTGGAGATGCTTAGTTT 360

QY 2168 TATTGCACAGAGTCAATTTGGGGCGTGCACCGCTGTTCTGGTATTCATTTTCATCA 2227
DB 361 TATTGCACAGAGTCAATTTGGGGCGTGCACCGCTGTTCTGGTATTCATTTTCATCA 420

QY 2228 CTGAGCATTTGATCATGCTTTTGGGCTTCTCAGTTCAATGAGCAATTAATGAAGTA 2287
DB 421 CTGAGCATTTGATCATGCTTTTGGGCTTCTCAGTTCAATGAGCAATTAATGAAGTA 480

QY 2288 TTAACTCTTTCACTACAGTCTTGTGCAAGTATGCTATTTAAATTTACTTGGCAGGTATA 2347
DB 481 TTAACTCTTTCACTACAGTCTTGTGCAAGTATGCTATTTAAATTTACTTGGCAGGTATA 540

QY 2348 TTGCCAGTCAGTCTCTTTATATAGTGAGAAAAATTTATGTTAGTATATAAATATTTTAAA 2407
DB 541 TTGCCAGTCAGTCTCTTTATATAGTGAGAAAAATTTATGTTAGTATATAAATATTTTAAA 600

QY 2408 CTAATATATAAATCTATAATGTTTAAACATATGTTTCAATTAAGACATAGCATTGAAAT 2467
DB 601 CTAATATATAAATCTATAATGTTTAAACATATGTTTCAATTAAGACATAGCATTGAAAT 660

QY 2468 TAACTATATAAATAGTCTCATATTTACACTTACAGCTTTTTCATTTGATCAGGTCCTGAAATC 2527
DB 661 TAACTATATAAATAGTCTCATATTTACACTTACAGCTTTTTCATTTGATCAGGTCCTGAAATC 720

QY 2528 TTTAGCACTTAAGGAAAAATGATGCTATGCTAATATATACCTGACCATGAAAAAATAGTAC 2587
DB 721 TTTAGCACTTAAGGAAAAATGATGCTATGCTAATATATACCTGACCATGAAAAAATAGTAC 780

QY 2588 CTCGAATCATGCTTTTGCACCT-GGTGATTCCTCACTGCACAAATCTTT 2634
DB 781 CTCGAATCATGCTTTTGCACCTGGGTGATTCCTCACTGCACAAATCTTT 828

RESULT 13
BUI45581
LOCUS BUI45581
DEFINITION AGENCOURT 8125917 Lupski_dorsal_root ganglion Homo sapiens cDNA
clone IMAGE:6177333 5', mRNA sequence.
ACCESSION BUI45581
VERSION BUI45581.1 GI:22659113
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 948)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. James R. Lupski
Email: cgabbs-remail.nih.gov
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13555 row: b column: 22
High quality sequence stop: 653.
Location/Qualifiers
1. 948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177333"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCCG-3' and
5'-GACTAGTCTAGATCGGAGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

FEATURES
source

ORIGIN
Query Match 25.4%; Score 815.2; DB 5; Length 948;
Query Match

Best Local Similarity 96.6%; Pred. No. 3e-178; Matches 887; Conservative 0; Mismatches 23; Indels 8; Gaps 5;	
Qy	1430 AAAATTCCTTTTCAGAAAATCAGCCATTTCTTCTTAAGATGCTTTTAGAGCTTTTGC 1489
Db	1 AAAATTCCTTTTCAGAAAATCAGCCATTTCTTCTTAAGATGCTTTTAGAGCTTTTGC 60
Qy	1490 CAGAAGATGATGGTTTAAACATTCGAATAAATGGAATCGCCACCAAGGGATGGAATGT 1549
Db	61 CAGAAGATGATGGTTTAAACATTCGAATAAATGGAATCGCCACCAAGGGATGGAATGT 120
Qy	1550 GGGATGGTAACTTATCAACATATTTTGACATGAATCTGTCTTTTGGATATATTTTAAAAA 1609
Db	121 GGGATGGTAACTTATCAACATATTTTGACATGAATCTGTCTTTTGGATATATTTTAAAAA 180
Qy	1610 CTGTTTTHAGAAAATCTCGGGAAGAGGAGAAATAGTGTCTTTTCTTCAATTTGATGCAGATATTT 1669
Db	181 CTGTTTTHAGAAAATCTCGGGAAGAGGAGAAATAGTGTCTTTTCTTCAATTTGATGCAGATATTT 240
Qy	1670 GCACATGGTTTCGGCAAAAGCAGACAAATATCCGATACTATTTTAACTCAAGGAAAT 1729
Db	241 GCACATGGTTTCGGCAAAAGCAGACAAATATCCGATACTATTTTAACTCAAGGAAAT 300
Qy	1730 CTGAGATTATCTGAACTCTGACCTCAGATCTCGGACAAACCCCAATTCGAATGAGCT 1789
Db	301 CTGAGATTATCTGAACTCTGACCTCAGATCTCGGACAAACCCCAATTCGAATGAGCT 360
Qy	1790 TTGCACAGTTTGAAAATCTACTGGGGATAAATGTACATATCTGAAGACTTGTCTCAGAAACC 1849
Db	361 TTGCACAGTTTGAAAATCTACTGGGGATAAATGTACATATCTGAAGACTTGTCTCAGAAACC 420
Qy	1850 CATCTTATATTCAGAGGCAAAAGCTAAGGAGCTAGTCAATATCTGCTGGGGTATGATA 1909
Db	421 CATCTTATATTCAGAGGCAAAAGCTAAGGAGCTAGTCAATATCTGCTGGGGTATGATA 480
Qy	1910 CCAATGATCTTGAACAGAGGAAATGAAGAACTTGGAGTAAATGCTCTAATTTATG 1969
Db	481 CCAATGATCTTGAACAGAGGAAATGAAGAACTTGGAGTAAATGCTCTAATTTATG 540
Qy	1970 ATAGGATATGATTTGGATGCTGAAACCAACCAATATATTCCAAGTGGAGCAATTTGGAAC 2029
Db	541 ATAGGATATGATTTGGATGCTGAAACCAACCAATATATTCCAAGTGGAGCAATTTGGAAC 600
Qy	2030 GCCTGAAGCAGAAATGCGAGCTTAAAGAGCTGTTGTGTGCCACTGTGACCGCTTTG 2089
Db	601 GCCTGAAGCAGAAATGCGAGCTTAAAGAGCTGTTGTGTGCCACTGTGACCGCTTTG 660
Qy	2090 TTCCCTCATCTTTGCTGGGAGTCTGATATCCATGCGATGCGCAACGGCAATTCATAACG 2149
Db	661 TTCCCTCATCTTTGCTGGGAGTCTGATATCCATGCGATGCGCAACGGCAATTCATAACG 720
Qy	2150 TGGAGATGCTTGTGTTTATTTGACAGAGGTCAATTTGGGGGGCTGCACCG-CTGTTCCT 2208
Db	721 TGGAGATGCTTGTGTTTATTTGACAGAGGTCAATTTGGGGGGCTGCACCGCTGTTCCT 780
Qy	2209 GGGTATTCATTTTTCATCACTGAGCAATTTGTTG---ATCTATGCTCTTTTGGGCTTCTCAGT 2265
Db	781 GGGTATTCATTTTTCATCACTGAGCAATTTGTTG---ATCTATGCTCTTTTGGGCTTCTCAGT 840
Qy	2266 TCAATGAAGCAAT--AATGAAGTATTTAACT-CTTTCACTACAG-TTCTTCAAGATATGC 2321
Db	841 TTCATGAAGCAATTAATGGAGGTATTTAACTCTCTTTCATCAGTTCTTTCGAAGATGTC 900
Qy	2322 TATTTTAAATTAATCTTGCC 2339
Db	901 TAAITAAAATACTTTGGC 918

RESULT 14
CR623938
LOCUS
DEFINITION
CR623938 811 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DB002YJ01 of Neuroblastoma Cot
10-normalized of Homo sapiens (human).

ACCESSION CR623938.1 GI:50504745 VERSION HTC; CNSLT_cDNA. KEYWORDS SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 811) Li W.B., Gruber C., Jessee J. and Polayes D. Full-length cDNA libraries and normalization Unpublished AUTHORS JOURNAL REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue 2 (bases 1 to 811) Genoscope. Direct Submission TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. FEATURES Location/Qualifiers source 1..811 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DB002YJ01" /tissue_type="Neuroblastoma Cot 10-normalized" /plasmid="pCMVSPORT_6" ORIGIN Query Match 25.2%; Score 809.4; DB 3; Length 811; Best Local Similarity 99.9%; Pred. No. 6.6e-177; Matches 810; Conservative 0; Mismatches 1; Indels 0; Gaps 0; Qy 2372 AGAAAATTTATGTTAGTAAATAATAATATTTTAAACTAAATATATAAAATCTATAATGTT 2431 Db 1 AGAAAATTTATGTTAGTAAATAATAATATTTTAAACTAAATATATAAAATCTATAATGTT 60 Qy 2432 AAACATATGTTTCAATTAAGCATAGACCTTTGAAATTAACATATATAAATAGTCATATTT 2491 Db 61 AAACATATGTTTCAATTAAGCATAGACCTTTGAAATTAACATATATAAATAGTCATATTT 120 Qy 2492 ACATTTACAGCTTTTCAATTTGATCAGGCTCTGAAATCTTTAGCACCTTAAGGAAATGACTA 2551 Db 121 ACATTTACAGCTTTTCAATTTGATCAGGCTCTGAAATCTTTAGCACCTTAAGGAAATGACTA 180 Qy 2552 TGCATAATTAACCTGCACCAAGGAAATTAAGTACCTCAAAATGCAATTTGCACTGG 2611 Db 181 TGCATAATTAACCTGCACCAAGGAAATTAAGTACCTCAAAATGCAATTTGCACTGG 240 Qy 2612 TGATTCCAACTGCACAAATCTTTTGCCACTCTTGATATATAGGTATTTTACATGGGTTG 2671 Db 241 TGATTCCAACTGCACAAATCTTTTGCCACTCTTGATATATAGGTATTTTACATGGGTTG 300 Qy 2672 ACATGCACAAACACCACTTTTCAATTCAGTATGAACCTTTGAGGCTGCTGCCATTTTCCAC 2731 Db 301 ACATGCACAAACACCACTTTTCAATTCAGTATGAACCTTTGAGGCTGCTGCCATTTTCCAC 360 Qy 2732 TTAACCAAAACAGCCTGAAGGTGAACCTCGAAAACCTTGTTTTCATAAAATCTTTCAAAAGTTG 2791 Db 361 TTAACCAAAACAGCCTGAAGGTGAACCTCGAAAACCTTGTTTTCATAAAATCTTTCAAAAGTTG 420 Qy 2792 TTTTACATCAATGTTAAAAATTTCAAAATGCTCGAGGGTAAATTTAATGTATAAAATATTAG 2851 Db 421 TTTTACATCAATGTTAAAAATTTCAAAATGCTCGAGGGTAAATTTAATGTATAAAATATTAG 480 Qy 2852 TAAGAAAAGTATGTTATGTCATCTTAGTAAATAGATCAACATACAAATTCATTTCA 2911 Db 480 TAAGAAAAGTATGTTATGTCATCTTAGTAAATAGATCAACATACAAATTCATTTCA 2911	
--	--

```
Db 481 TAAGAAAAGTATGTAATGTCATCTTAGTAGAATAGATCACAACATCAAAATTCATTTCA 540
Qy 2912 GTGCATGCTTTAGGTGTTAAAGCATGAGATTTGATCATGTTTACTGTAGGTCTTGCATCT 2971
Db 541 GTGCATGCTTTAGGTGTTAAAGCATGAGATTTGATCATGTTTACTGTAGGTCTTGCATCT 600
Qy 2972 GTGGTGCTAGGTGATGAGAGATGTCAGGACTGACGATGATTTGTTGGCTCAAAAAA 3031
Db 601 GTGGTGCTAGGTGATGAGAGATGTCAGGACTGACGATGATTTGTTGGCTCAAAAAA 660
Qy 3032 AAAAGCGTGTGTTAGGCGTTTAAATATGCTTATTTGTTGTGCTCTCACCTACCTATTA 3091
Db 661 AAAAGCGTGTGTTAGGCGTTTAAATATGCTTATTTGTTGTGCTCTCACCTACCTATTA 720
Qy 3092 CACACTGTGCTTTGTGGGTTGTTGTTGATGCGGTGTTGTTATACAGTAGTTAAATTTTC 3151
Db 721 CACACTGTGCTTTGTGGGTTGTTGTTGATGCGGTGTTGTTATACAGTAGTTAAATTTTC 780
Qy 3152 CATGCAGAAAAATAAATGCTCCTGAATTTTCA 3182
Db 781 CATGCAGAAAAATAAATGCTCCTGAATTTTCA 811

RESULT 15
AL520917 811 bp mRNA linear EST 24-MAR-2004
LOCUS AL520917 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DB002VJ01 5-PRIME, mRNA sequence.
ACCESSION AL520917
VERSION AL520917.3 GI:45696403
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:31039233.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9838.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DB002CE01QP1&c=9838.r.
FEATURES
source
1..811
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB002VJ01"
/tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 24.9%; Score 799.8; DB 1; Length 811;
Best Local Similarity 98.8%; Pred. No. 1.1e-174;
Matches 801; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2372 AGAAAATTTATGGTTAGTAATAATAATTTTAACTAAATATATAATCTATTAATGTT 2431
Db 1 AGAAAATTTATGGTTAGTAATAATAATTTTAACTAAATATATAATCTATTAATGTT 60
```

```
Qy 2432 AAACATATGTTCAATAAAGCATAGCACTTTGAAATTAACATATATAAATAGCTCATATTT 2491
Db 61 AAACATATGTTCAATAAAGCATAGCACTTTGAAATTAACATATATAAATAGCTCATATTT 120
Qy 2492 ACACCTACAGCTTTTCATTCATGAGGTCTGAAATCTTTAGCACTTAGGAAAATGACTA 2551
Db 121 ACACCTACAGCTTTTCATTCATGAGGTCTGAAATCTTTAGCACTTAGGAAAATGACTA 180
Qy 2552 TGCATAATATATACCTGACCATGAAAAAATAAGTACCTCAATGATGATGCAATTTGCACCTGG 2611
Db 181 TGCATAATATATACCTGACCATGAAAAAATAAGTACCTCAATGATGATGCAATTTGCACCTGG 240
Qy 2612 TGATTCCAACTGCACAAATCTTTGTGCCATCTTGTATATAGGTATTTTTCATCGGTTG 2671
Db 241 TGATTCCAACTGCACAAATCTTTGTGCCATCTTGTATATAGGTATTTTTCATCGGTTG 300
Qy 2672 ACATGCACACAAACCACTTTTCATTCAGTATGAACCTTGAGGCTGCGCCATTTTCCAC 2731
Db 301 ACATGCACACAAACCACTTTTCATTCAGTATGAACCTTGAGGCTGCGCCATTTTCCAC 360
Qy 2732 TTAACCAAAACAGCTGAAAGGTGAACCTCGAAACCTTGTTTCATAAATCTTTCAAAAGTTG 2791
Db 361 TTAACCAAAACAGCTGAAAGGTGAACCTCGAAACCTTGTTTCATAAATCTTTCAAAAGTTG 420
Qy 2792 TTTTACATCAATGTTTAAATTTTCAAAATGCTGCGGGTAAATTTAATGTATATAAATATTAG 2851
Db 421 TTTTACATCAATGTTTAAATTTTCAAAATGCTGCGGGTAAATTTAATGTATATAAATATTAG 480
Qy 2852 TAAGAAAAAGTATGTTATGTCATCTTAGTAGAATAGATCACAACATCAAAATTCATTTCA 2911
Db 481 TAAGAAAAAGTATGTTATGTCATCTTAGTAGAATAGATCACAACATCAAAATTCATTTCA 540
Qy 2912 GTGCATGCTTTAGGTGTTTAAAGCATGAGATTTGATCATGTTTACTGTAGGTCTTGCATCT 2971
Db 541 GTGCATGCTTTAGGTGTTTAAAGCATGAGATTTGATCATGTTTACTGTAGGTCTTGCATCT 600
Qy 2972 GTGGTGCTAGGTGATGAGAGATGTCAGGACTGACGATGATTTGTTGGCTCAAAAAA 3031
Db 601 GTGGTGCTAGGTGATGAGAGATGTCAGGACTGACGATGATTTGTTGGCTCAAAAAA 660
Qy 3032 AAAAGCGTGTGTTAGGCGTTTAAATATGCTTATTTGTTGTGCTCTCACCTACCTATTA 3091
Db 661 AAAAGCGTGTGTTAGGCGTTTAAATATGCTTATTTGTTGTGCTCTCACCTACCTATTA 720
Qy 3092 CACACTGTGCTTTGTGGGTTGTTGTTGATGCGGTGTTGTTATACAGTAGTTAAATTTTC 3151
Db 721 CACACTGTGCTTTGTGGGTTGTTGTTGATGCGGTGTTGTTATACAGTAGTTAAATTTTC 780
Qy 3152 CATGCAGAAAAATAAATGCTCCTGAATTTTCA 3182
Db 781 CATGCAGAAAAATAAATGCTCCTGAATTTTCA 811
```

Search completed: July 2, 2005, 19:42:19
Job time : 9990 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 12:38:55 ; Search time 13649 Seconds
(without alignments)
11381.597 Million cell updates/sec

Title: US-10-047-855-4
Perfect score: 3206
Sequence: 1 gtcgaccacgcgtccggc.....aaaaaaaaggcgccgcg 3206

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3206	100.0	3206	6	AX127528 Sequence
2	3206	100.0	3206	6	AX675186 Sequence
3	3178	99.1	3265	9	BC027588 Homo sapi
4	3166	98.8	3499	6	CO413929 Sequence
5	3166	98.8	5443	9	AB037855 Homo sapi
6	3164.4	98.7	3499	6	CQ491863 Sequence
7	3164.4	98.7	3499	6	CQ497765 Sequence
8	2726.8	85.1	6671	9	HSN804380
9	2481.8	77.4	2738	6	AX127525 Sequence
10	2329.4	72.7	39694	2	AC144439 Homo sapi
11	2329.4	72.7	149259	9	AC099654 Homo sapi
12	2292.6	71.5	160762	2	AC150686 Pan trogl
13	1897.6	59.2	3229	10	AV233980 Rattus no
14	1803	56.2	1803	6	BD156897 Primer fo
15	1803	56.2	1803	6	AX877873 Sequence
16	1803	56.2	1803	9	AK001947 Homo sapi
17	1645.2	51.3	3174	10	BC033408 Homo sapi
18	1462	45.6	1731	6	CQ728047 Sequence
19	1369.4	42.7	1597	11	BV177434 sqnm94654

C	20	1369.4	42.7	1597	11	BV180479	BV180479 sqnm10972
C	21	1215.2	37.9	178601	9	HSJ1022P6	AL109935 Human DNA
C	22	1049.6	32.7	4353	10	AK122510	AK122510 Mus muscu
C	23	633.4	19.8	768	6	BD149663	BD149663 Primer fo
C	24	633.4	19.8	768	6	AX869601	AX869601 Sequence
C	25	606.8	18.9	85916	2	AC023409	AC023409 Homo sapi
C	26	542.4	16.9	183076	10	AL807386	AL807386 Mouse DNA
C	27	535.4	16.7	245453	2	AC131004	AC131004 Rattus no
C	28	535.4	16.7	252563	2	AC094054	AC094054 Rattus no
C	29	486.4	15.2	546	6	AX197615	AX197615 Sequence
C	30	486.4	15.2	546	6	AX208231	AX208231 Sequence
C	31	444.6	13.9	490	6	BD154638	BD154638 Primer fo
C	32	444.6	13.9	490	6	AX874576	AX874576 Sequence
C	33	438.2	13.7	273082	2	AC128177	AC128177 Rattus no
C	34	414.8	12.9	520	6	CQ409598	CQ409598 Sequence
C	35	413.8	12.9	610	9	HSN800231	AL049446 Homo sapi
C	36	365.6	11.4	403	6	AX197736	AX197736 Sequence
C	37	365.6	11.4	403	6	AX208343	AX208343 Sequence
C	38	363.4	11.3	396	6	CQ396906	CQ396906 Sequence
C	39	363.4	11.3	396	6	CQ403214	CQ403214 Sequence
C	40	326.6	10.2	425	6	CQ504016	CQ504016 Sequence
C	41	326.6	10.2	425	6	CQ513079	CQ513079 Sequence
C	42	326	10.2	327	6	CQ483187	CQ483187 Sequence
C	43	296.4	9.2	478	6	CQ461585	CQ461585 Sequence
C	44	291	9.1	510	6	CQ459591	CQ459591 Sequence
C	45	271	8.5	359	6	CQ474018	CQ474018 Sequence

ALIGNMENTS

RESULT 1
AX127528 3206 bp DNA linear PAT 15-MAY-2001
LOCUS AX127528 Sequence 4 from Patent WO0131007.
DEFINITION AX127528
ACCESSION AX127528.1 GI:14134233
VERSION
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chiang, L.W.
TITLE Nucleic acid molecules derived from rat brain and programmed cell death models
JOURNAL Patent: WO 0131007-A 4 03-MAY-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1..3206
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 3206; DB 6; Length 3206;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTCGACCCACCGCTCCGGCGGCGGAGGACGAGCGGGCGCGCGGTACCTCTGCGCGCGGT 60
Db 1 GTCGACCCACCGCTCCGGCGGCGGAGGACGAGCGGGCGCGCGGTACCTCTGCGCGCGGT 60
Qy 61 CCTCGCTCTCGGGCGGCGGCGGCGGAGCTCGCGGACTAGCGAACCCGGAGCAGC 120
Db 61 CCTCGCTCTCGGGCGGCGGCGGCGGAGCTCGCGGACTAGCGAACCCGGAGCAGC 120
Qy 121 ACATCAATAAATAATCCATCAGATGACACCTTCTCAGGTTCGCTTTGAATAAGAGGA 180
Db 121 ACATCAATAAATAATCCATCAGATGACACCTTCTCAGGTTCGCTTTGAATAAGAGGA 180
Qy 181 ACTCTTTTACAGGAGAGGATTTTTCGCGATATGTGGAAGCTGTGATGCTTTCGGAACTGG 240
Db 181 ACTCTTTTACAGGAGAGGATTTTTCGCGATATGTGGAAGCTGTGATGCTTTCGGAACTGG 240

Db 181 ACTCTTTTACCAGGAGAGTTTTTGGCGATATGTGGAAGCTGTGATGCTTTGGGAAACTGG 240
Qy 241 AATCCTCAAAATGCTGTGGCTCTCTTCCAGAGAAATGACACAGGTGAAGCATGCTATGG 300
Db 241 AATCCTCAAAATGCTGTGGCTCTCTTCCAGAGAAATGACACAGGTGAAGCATGCTATGG 300
Qy 301 AAAGCAACCATTTACTCAGTAGAGGAGTATCAGTTCAGTATCGCTACTCTCAAGGGTAC 360
Db 301 AAAGCAACCATTTACTCAGTAGAGGAGTATCAGTTCAGTATCGCTACTCTCAAGGGTAC 360
Qy 361 TTTTGTAGAACCAAGACATATCGGTGGTCCATGTCAAGTGTAGTTTCAAAAGTGGAGACT 420
Db 361 TTTTGTAGAACCAAGACATATCGGTGGTCCATGTCAAGTGTAGTTTCAAAAGTGGAGACT 420
Qy 421 CATCTAACCAACGATCAATAACCCCTTTAGAAAGCGAAATTAATTTAGCATGTGACAA 480
Db 421 CATCTAACCAACGATCAATAACCCCTTTAGAAAGCGAAATTAATTTAGCATGTGACAA 480
Qy 481 TTTTGGAAATCCAAATGTTGTAAGTCTGAACTCTGAACTCTGAACTCTGAACTCTGAA 540
Db 481 TTTTGGAAATCCAAATGTTGTAAGTCTGAACTCTGAACTCTGAACTCTGAACTCTGAA 540
Qy 541 ATAAGATTACGTTTGCATTAATCTGAAATACTCTGTGTCATTAACCAAGAAAAATTA 600
Db 541 ATAAGATTACGTTTGCATTAATCTGAAATACTCTGTGTCATTAACCAAGAAAAATTA 600
Qy 601 AAAAAATCTAGATTAGGGTGAAGCTGACACTAGAGGCCCTGGAGGAAGATGACGATGAT 660
Db 601 AAAAAATCTAGATTAGGGTGAAGCTGACACTAGAGGCCCTGGAGGAAGATGACGATGAT 660
Qy 661 AGGGTATCTCCACCTGTACTCCACAAAATGTCCATAGCTTTGGAGATATCCCTTAATAGC 720
Db 661 AGGGTATCTCCACCTGTACTCCACAAAATGTCCATAGCTTTGGAGATATCCCTTAATAGC 720
Qy 721 GACAAATGAGTTCAAGTGCAGGCAATTCACAGCGAGTGTGTTATGGCTTCAGCCTGAT 780
Db 721 GACAAATGAGTTCAAGTGCAGGCAATTCACAGCGAGTGTGTTATGGCTTCAGCCTGAT 780
Qy 781 CGTTGGACAGAGTACAGCATACAGATGGAACACAGATAACCTGGAACTAATCTTTGAT 840
Db 781 CGTTGGACAGAGTACAGCATACAGATGGAACACAGATAACCTGGAACTAATCTTTGAT 840
Qy 841 TTTTTCGAAGAAGATCTCAGTGAGCAGTGTTCAGGGTGTATGCCCTTCCTGGACATGTG 900
Db 841 TTTTTCGAAGAAGATCTCAGTGAGCAGTGTTCAGGGTGTATGCCCTTCCTGGACATGTG 900
Qy 901 GGTACAGCTTGTCTCTTATCATCCACCATGTCTGAGAGTGGAAAGAGTGTGGAAATCTT 960
Db 901 GGTACAGCTTGTCTCTTATCATCCACCATGTCTGAGAGTGGAAAGAGTGTGGAAATCTT 960
Qy 961 ACTCTTCCCATCATGAGCAGAAATTTCCCGGAAAAACAATAGGCAAAAGTGAGAGTTGACTAT 1020
Db 961 ACTCTTCCCATCATGAGCAGAAATTTCCCGGAAAAACAATAGGCAAAAGTGAGAGTTGACTAT 1020
Qy 1021 ATAAATTTAAGCCATTAACAGGATACAGTTGTGATGATGAATCTTCAATTTTCCAGTAT 1080
Db 1021 ATAAATTTAAGCCATTAACAGGATACAGTTGTGATGATGAATCTTCAATTTTCCAGTAT 1080
Qy 1081 TGAAGCCCAAGAAATACCATTTGATGTGGCATCGAGTGCAGGAACTCTTACACAACT 1140
Db 1081 TGAAGCCCAAGAAATACCATTTGATGTGGCATCGAGTGCAGGAACTCTTACACAACT 1140
Qy 1141 GCCAGCTGGCTAAAGTTCAAGAAATACTATTGCTTCTTTAAGAAATGCTGTAGTCAAT 1200
Db 1141 GCCAGCTGGCTAAAGTTCAAGAAATACTATTGCTTCTTTAAGAAATGCTGTAGTCAAT 1200
Qy 1201 GGTGACGCTTTGTAGAAATTTGACGTACACCTTTCAAAGGACTTTGTGCCCGGTGTATAT 1260
Db 1201 GGTGACGCTTTGTAGAAATTTGACGTACACCTTTCAAAGGACTTTGTGCCCGGTGTATAT 1260
Qy 1261 CATGATCTTACCTGTTGTTGACTATGAAGAAATTTGATGCTGATCCAGTTCGAATTA 1320
Db 1261 CATGATCTTACCTGTTGTTGACTATGAAGAAATTTGATGCTGATCCAGTTCGAATTA 1320

Qy 1321 TTTGAAATTTCCAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCAT 1380
Db 1321 TTTGAAATTTCCAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCAT 1380
Qy 1381 GTGACTGCATCGAAATCTAAGGATCGGAAAGAAATCTGTGGTTTCAGAGGAAAAATTCCTTT 1440
Db 1381 GTGACTGCATCGAAATCTAAGGATCGGAAAGAAATCTGTGGTTTCAGAGGAAAAATTCCTTT 1440
Qy 1441 TCAGAAATCAGCCATTTCTCTCTTAAGATGTTTAAAGTCTTTAGAGTCTTTGCGCAAGATGTA 1500
Db 1441 TCAGAAATCAGCCATTTCTCTCTTAAGATGTTTAAAGTCTTTAGAGTCTTTGCGCAAGATGTA 1500
Qy 1501 GGGTTTAAATTAATAAAATGGAATCTGCGACGAAAGGATGGAATGTGGATGGTAAC 1560
Db 1501 GGGTTTAAATTAATAAAATGGAATCTGCGACGAAAGGATGGAATGTGGATGGTAAC 1560
Qy 1561 TTATCAACATATTTTGGACATGAATCTGTTTGGATATATATTTTAAAAAATCTGTTTAGAA 1620
Db 1561 TTATCAACATATTTTGGACATGAATCTGTTTGGATATATATTTTAAAAAATCTGTTTAGAA 1620
Qy 1621 AATTTCTGGGAAGAGGAGAAATAGTGTCTTCAATTTTGAATGTCAGATATTTGCAAAATGGTT 1680
Db 1621 AATTTCTGGGAAGAGGAGAAATAGTGTCTTCAATTTTGAATGTCAGATATTTGCAAAATGGTT 1680
Qy 1681 CGGCAAAAGCAGAAATAATCCGATACTATTTTAACTCAAGGAAAAATCTGAGATTTAT 1740
Db 1681 CGGCAAAAGCAGAAATAATCCGATACTATTTTAACTCAAGGAAAAATCTGAGATTTAT 1740
Qy 1741 CCTGAACTCATGGACCTCAGATCTCGGACAAACCCCAATGCAATGAGCTTTGCACAGTTT 1800
Db 1741 CCTGAACTCATGGACCTCAGATCTCGGACAAACCCCAATGCAATGAGCTTTGCACAGTTT 1800
Qy 1801 GAAAACTCTAATGGGATATAATGTATGATGAAGACTTCTCAGAAACCCATCTATAT 1860
Db 1801 GAAAACTCTAATGGGATATAATGTATGATGAAGACTTCTCAGAAACCCATCTATAT 1860
Qy 1861 CAGAGGCAAAAGCTAAGGAGCTAGTATATTTCTGCTGGGTGATGATACCAATGATCT 1920
Db 1861 CAGAGGCAAAAGCTAAGGAGCTAGTATATTTCTGCTGGGTGATGATACCAATGATCT 1920
Qy 1921 GAAAAACAAGGAAATTTGAAGGAACTTTGGAGTTAATGGTCTAATTTATGATAGGATATAT 1980
Db 1921 GAAAAACAAGGAAATTTGAAGGAACTTTGGAGTTAATGGTCTAATTTATGATAGGATATAT 1980
Qy 1981 GATTTGATGCTGAAACAAACCAATATATTTCCAAAGTGGAGCAATTTGGAACCCCTGAAAGCAG 2040
Db 1981 GATTTGATGCTGAAACCAACCAATATATTTCCAAAGTGGAGCAATTTGGAACCCCTGAAAGCAG 2040
Qy 2041 GAATTTGCCAGAGCTTAAGAGCTGTTGTGTCACACTGTTAGCCGCTTTGTCCTCATCT 2100
Db 2041 GAATTTGCCAGAGCTTAAGAGCTGTTGTGTCACACTGTTAGCCGCTTTGTCCTCATCT 2100
Qy 2101 TTGTGTGGGAGTCTGATATCCATGTGGATGCCAACGGCAATTTGATAACGTGGAGAAATGCT 2160
Db 2101 TTGTGTGGGAGTCTGATATCCATGTGGATGCCAACGGCAATTTGATAACGTGGAGAAATGCT 2160
Qy 2161 TAGTTTTTATTTGACACAGAGGTCAATTTTGGGGGCGTGACCGCTGTTCTGGGTATTCATTT 2220
Db 2161 TAGTTTTTATTTGACACAGAGGTCAATTTTGGGGGCGTGACCGCTGTTCTGGGTATTCATTT 2220
Qy 2221 TTGATCTCAGAGCAATTTGTTGATCTATGCTTTTGGGCTTCTCAGTTCAATGAAGCAATAA 2280
Db 2221 TTGATCTCAGAGCAATTTGTTGATCTATGCTTTTGGGCTTCTCAGTTCAATGAAGCAATAA 2280
Qy 2281 TGAAGTATTTAACTCTTTTCACTACAGTCTTGCAAGTATGCTATTTAAATTTACTTGGSCCA 2340
Db 2281 TGAAGTATTTAACTCTTTTCACTACAGTCTTGCAAGTATGCTATTTAAATTTACTTGGSCCA 2340
Qy 2341 GGTATAATTTGCCAGTCACTCTTTATAGTGAGAAATTTTATGTTAGTAAATATAAATA 2400
Db 2341 GGTATAATTTGCCAGTCACTCTTTATAGTGAGAAATTTTATGTTAGTAAATATAAATA 2400

QY	2401	TTTTAACTAAATATATAATCTATTAATGTTAAACATATGTTCAATTAAGCATAGCACT	2460
Db	2401	TTTTAACTAAATATATAATCTATTAATGTTAAACATATGTTCAATTAAGCATAGCACT	2460
QY	2461	TTGAAATTAACATATAAATAGCTCATATTTACACATTACAGCTTTTTCATTTGATCAGGTC	2520
Db	2461	TTGAAATTAACATATAAATAGCTCATATTTACACATTACAGCTTTTTCATTTGATCAGGTC	2520
QY	2521	TGAAATCTTTTAGCATTAAAGGAAATGACTATGTCATAATATACCTGACCATGAAAAAA	2580
Db	2521	TGAAATCTTTTAGCATTAAAGGAAATGACTATGTCATAATATACCTGACCATGAAAAAA	2580
QY	2581	TAAGTACCTCAATAGCATGATTTGACATGCTGTTGATTCACACTGCACAAAATCTTTGTGCCA	2640
Db	2581	TAAGTACCTCAATAGCATGATTTGACATGCTGTTGATTCACACTGCACAAAATCTTTGTGCCA	2640
QY	2641	TCCTGTATATAGTATTTTTTACATGGTGTGACATGCACACACACCAATTTTCATTCAGT	2700
Db	2641	TCCTGTATATAGTATTTTTTACATGGTGTGACATGCACACACACCAATTTTCATTCAGT	2700
QY	2701	ATGAACCTTTGAGGCTGCTGCCATTTTCCACTTAACCAACCAAGCTGGAAGGTGAACCTC	2760
Db	2701	ATGAACCTTTGAGGCTGCTGCCATTTTCCACTTAACCAACCAAGCTGGAAGGTGAACCTC	2760
QY	2761	GAAACTGTTTTCATAAATCTTTCAAAAGTTGTTTACATCAATGTTTAAATTTCAAAATG	2820
Db	2761	GAAACTGTTTTCATAAATCTTTCAAAAGTTGTTTACATCAATGTTTAAATTTCAAAATG	2820
QY	2821	CTGCGGGTAATTTAATGTATATAATATTAGTAAGAAAGTATGTTGTCATACCTAGT	2880
Db	2821	CTGCGGGTAATTTAATGTATATAATATTAGTAAGAAAGTATGTTGTCATACCTAGT	2880
QY	2881	AGAATAGATCAACAATACAAATTCAAATTCAGTGCATGCTTTAGTGTGTTAAGCATGAGAT	2940
Db	2881	AGAATAGATCAACAATACAAATTCAAATTCAGTGCATGCTTTAGTGTGTTAAGCATGAGAT	2940
QY	2941	TGTACATGTTTACTGTAGTCTGTCATCTGTTGCTGTTAGTGTAGTGTAGTGTAGTGTAGT	3000
Db	2941	TGTACATGTTTACTGTAGTCTGTCATCTGTTGCTGTTAGTGTAGTGTAGTGTAGTGTAGT	3000
QY	3001	RAAGGACTGACCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3060
Db	3001	RAAGGACTGACCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3060
QY	3061	GCTTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3120
Db	3061	GCTTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3120
QY	3121	ATGTCGGTGTATACAGTATGTTAAATTTCCATGTCAGAAAAATAAATGCTGTAATTC	3180
Db	3121	ATGTCGGTGTATACAGTATGTTAAATTTCCATGTCAGAAAAATAAATGCTGTAATTC	3180
QY	3181	CAAAAAAAGGAGGCGCGC	3206
Db	3181	CAAAAAAAGGAGGCGCGC	3206
RESULT 2			
LOCUS	AX675186		
DEFINITION	Sequence 4 from Patent WO02081516.		
ACCESSION	AX675186		
VERSION	AX675186.1	GI:29333417	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1		
JOURNAL	Chiang, L.W.		
	Narc10 and narc16, programmed cell death-associated molecules and		
	uses thereof		
	Patent: WO 02081516-A 4 17-OCT-2002;		

FEATURES		Millennium Pharmaceuticals, Inc. (US)	
source		Location/Qualifiers	
		1..3206	
		/organism="Homo sapiens"	
		/mol_type="Unassigned DNA"	
		/db_xref="taxon:9606"	
misc_feature		1..17	
		/note="Vector sequence"	
CDS		145..2163	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD83808.1"	
		/db_xref="GI:29333418"	
		/translation="MTPSQVAFIRGTLIPGEFAICGSCDALGNWNPQNAVALPEN	
		DTGSMMLKATIVLSRGVQYRYFGYFLEPKTIGGCPQIVHKWETHLOPRSTPL	
		ESEIIDDQFGIHNGVETLDSGLTCTQTEIRLRHYSEKPPVSTTKKKLKRFRVK	
		LTLSEGLDEDDRVSPTVLHKSNSLEISLSDNEFKRHSQPEGGVGLQDRWTEYS	
		IQTEPNDLELIDFEEDESHVVGDALEPHVGTACLLSTTSAESKSGAGILFLPI	
		MSRNRKTIGKRVYDIILKPLPGYSCDKSSFSKYMKPRIPLCLVHKKGTADTAEQ	
		LAKVQENTIASLRNAASHGAAFVDFVHLSKDFVYVYHDLTCLTMKKKTDADVEL	
		FEIPVKELTDLQQLKLTHTALKSKDKESVVOEENSFSFENOPFPLKMLVLSLPE	
		DVGFNIEIKIQQRDGMWGNLSTYFDMNLFLDILKTVLENSGRRIIVSSFPADI	
		CTMVKONKYPILFLQCKSEIYPELMDLRSTPIAMSPAQENLGINVHVEDLL	
		RNPSYIQAKAGLVIFCWGDDTNDPNNRKLKELGVNGLIYDIYDMPSQPNIFQV	
		EQLERLKOELPELKLCLPTVSRFVPSLGSEDIHVDANGINDVNA"	
ORIGIN			
Query Match		100.0%; Score 3206; DB 6; Length 3206;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 3206; Conservative		0; Mismatches	
		0; Indels	
		0; Gaps	
QY	1	GTGACCCACACGCTCGGGCGAGGACGAGCGGGCGGGCGGGTACCTCTGCCCGCGT	60
Db	1	GTGACCCACACGCTCGGGCGAGGACGAGCGGGCGGGTACCTCTGCCCGCGT	60
QY	61	CCTCGCTCTCGGGCGGGCGGGCGAGCGGACCTCGGGACTAGCGAACCGGAGCAG	120
Db	61	CCTCGCTCTCGGGCGGGCGGGCGAGCGGACCTCGGGACTAGCGAACCGGAGCAG	120
QY	121	ACATCATATAAATAATCCATCAGATGACCTTCTCAGTTCGCTTTGAAATAAGAGA	180
Db	121	ACATCATATAAATAATCCATCAGATGACCTTCTCAGTTCGCTTTGAAATAAGAGA	180
QY	181	ACTCTTTTACAGGAGAGTCTTTCGATATGCGAGCTGTGATGCTTTGGGAACTGG	240
Db	181	ACTCTTTTACAGGAGAGTCTTTCGATATGCGAGCTGTGATGCTTTGGGAACTGG	240
QY	241	AATCTCTCAAAATGCTGTGGCTCTTCTCCAGAGATGACAGGTGAAAGCATGCTATGG	300
Db	241	AATCTCTCAAAATGCTGTGGCTCTTCTCCAGAGATGACAGGTGAAAGCATGCTATGG	300
QY	301	AAAGCAACCATTTGCTCAGTAGAGGAGTATCAGTTCAGTATCGCTACTTCAAAAGGTAC	360
Db	301	AAAGCAACCATTTGCTCAGTAGAGGAGTATCAGTTCAGTATCGCTACTTCAAAAGGTAC	360
QY	361	TTTTTAGAACCAAGACTATCGGTGCTCCATGTCAGTGTAGTGTACAGTGGGAGACT	420
Db	361	TTTTTAGAACCAAGACTATCGGTGCTCCATGTCAGTGTAGTGTACAGTGGGAGACT	420
QY	421	CATCTACCAACCATCAATTAACCCCTTTTAGAAGCGGAAATATTATTGACGATGGACAA	480
Db	421	CATCTACCAACCATCAATTAACCCCTTTTAGAAGCGGAAATATTATTGACGATGGACAA	480
QY	481	TTTGGAAATCCAAATGTTGTTGAAATCTGATTCGATTCGCTGATGTCAGACTGAA	540
Db	481	TTTGGAAATCCAAATGTTGTTGAAATCTGATTCGATTCGCTGATGTCAGACTGAA	540
QY	541	ATAAGATTACGTTTGTGCAATTTATTTGAAAAAACCCTCTGTGTCAATAACCAAGAAAAATTA	600
Db	541	ATAAGATTACGTTTGTGCAATTTATTTGAAAAAACCCTCTGTGTCAATAACCAAGAAAAATTA	600
QY	601	AAAAAATCTAGATTAGGTTGAACTGACACTAGAACGCTTGGAGGAAGATGACGATGAT	660
Db	601	AAAAAATCTAGATTAGGTTGAACTGACACTAGAACGCTTGGAGGAAGATGACGATGAT	660

Db 601 AAAAAATCTAGATTAGGGTGAAGCTGACACTAGAAGGCGCTGGAGGAAGATGACGATGAT 660
Qy 661 AGGGTATCTCCCACTGTACTCCACAAAATGTCCAATAGCTTGGAGATATCTCTTAATAAGC 720
Db 661 AGGGTATCTCCCACTGTACTCCACAAAATGTCCAATAGCTTGGAGATATCTCTTAATAAGC 720
Qy 721 GACAAATGAGTTCAAGTGCAGGCATTTACAGCCGGAGTGTGGTTATGGCTTGCAGCCCTGAT 780
Db 721 GACAAATGAGTTCAAGTGCAGGCATTTCAAGCCGGAGTGTGGTTATGGCTTGCAGCCCTGAT 780
Qy 781 CGTTTGGACAGAGTACAGCATACAGACGATGGAACACAGATAACCTGGAACTAAATCTTTGAT 840
Db 781 CGTTTGGACAGAGTACAGCATACAGACGATGGAACACAGATAACCTGGAACTAAATCTTTGAT 840
Qy 841 TTTTTTCCGAAGAAGTCTCAGTGCAGCAGTAGTTTCAGGGTGATGCCCTTCTCGTGACATGTG 900
Db 841 TTTTTTCCGAAGAAGTCTCAGTGCAGCAGTAGTTTCAGGGTGATGCCCTTCTCGTGACATGTG 900
Qy 901 GGTACAGCTTGTCTCTTATCATCCACCATTGCTGAGAGTGGAAAGAGTGTGGAAATCTCT 960
Db 901 GGTACAGCTTGTCTCTTATCATCCACCATTGCTGAGAGTGGAAAGAGTGTGGAAATCTCT 960
Qy 961 ACTCTTCCCATCATGAGCAGAGAAATTCGCGGAAAAAATAAGCAAGTGCAGATTTGACTAT 1020
Db 961 ACTCTTCCCATCATGAGCAGAGAAATTCGCGGAAAAAATAAGCAAGTGCAGATTTGACTAT 1020
Qy 1021 ATAATTAATTAAGCCATTACCAAGATACAGTTGTGACATGAATACTTCAATTTTCCAAGTAT 1080
Db 1021 ATAATTAATTAAGCCATTACCAAGATACAGTTGTGACATGAATACTTCAATTTTCCAAGTAT 1080
Qy 1081 TGGAAAGCCAAAGTAACCAATTCAGTGTGGCCATCGAGGTGAGGAACTCTTACAACT 1140
Db 1081 TGGAAAGCCAAAGTAACCAATTCAGTGTGGCCATCGAGGTGAGGAACTCTTACAACT 1140
Qy 1141 GCCAGCTGGCTTAAAGTTCAAGAAATACTATGCTTCTTTAAAGAAATGCTGTAGTCAT 1200
Db 1141 GCCAGCTGGCTTAAAGTTCAAGAAATACTATGCTTCTTTAAAGAAATGCTGTAGTCAT 1200
Qy 1201 GGTGACGCTTTGTAGAAATTTGACGTACACCTTTTCAAGAGGACTTTGTGCCGCTGGTATAT 1260
Db 1201 GGTGACGCTTTGTAGAAATTTGACGTACACCTTTTCAAGAGGACTTTGTGCCGCTGGTATAT 1260
Qy 1261 CATGATCTTACCTGTGTGTGACTATGAAAGAAATTTGATGCTGATCCAGTTGAAATTA 1320
Db 1261 CATGATCTTACCTGTGTGTGACTATGAAAGAAATTTGATGCTGATCCAGTTGAAATTA 1320
Qy 1321 TTTGAAATTCAGTAAGAAATTAACATTTGACCAACTCCAGTTGCTTAAAGCTCACTCAT 1380
Db 1321 TTTGAAATTCAGTAAGAAATTAACATTTGACCAACTCCAGTTGCTTAAAGCTCACTCAT 1380
Qy 1381 GTGACTGCACTGAAATCTTAAGGATCGGAAAGAAATCTGTGGTTTCAGGAGGAAATTTCTTTT 1440
Db 1381 GTGACTGCACTGAAATCTTAAGGATCGGAAAGAAATCTGTGGTTTCAGGAGGAAATTTCTTTT 1440
Qy 1441 TCAGAAATTCAGCAATTCCTTCTCTTAAGATGGTTTATAGTCTTTTGCAGAGATGTA 1500
Db 1441 TCAGAAATTCAGCAATTCCTTCTCTTAAGATGGTTTATAGTCTTTTGCAGAGATGTA 1500
Qy 1501 GGGTTTAAATTTGAATAAATGGATCTGCCAGCAAGGATGGAAATGGGATGGTAAAC 1560
Db 1501 GGGTTTAAATTTGAATAAATGGATCTGCCAGCAAGGATGGAAATGGGATGGTAAAC 1560
Qy 1561 TTATCAACATATTTTGCATGAATCTGTTTTTGGATATAATTTTAAAACTGTTTTAGAA 1620
Db 1561 TTATCAACATATTTTGCATGAATCTGTTTTTGGATATAATTTTAAAACTGTTTTAGAA 1620
Qy 1621 AATTCCTGGGAAGAGGAAATAGTGTGTTTTCTTCAATTTGATGCGATATTTTGCACATGGTT 1680
Db 1621 AATTCCTGGGAAGAGGAAATAGTGTGTTTTCTTCAATTTGATGCGATATTTTGCACATGGTT 1680
Qy 1681 CGGCAAAAGCAGAAACAAATATCCGATACTATTTTAACTCAAGGAAATCTCGAGATTTAT 1740
Db 1681 CGGCAAAAGCAGAAACAAATATCCGATACTATTTTAACTCAAGGAAATCTCGAGATTTAT 1740

Qy 1741 CCTGAACCTCATGGACCTCAGATCTCGGACAAACCCCAATTTGCAATGAGCTTTTGCACAGTTT 1800
Db 1741 CCTGAACCTCATGGACCTCAGATCTCGGACAAACCCCAATTTGCAATGAGCTTTTGCACAGTTT 1800
Qy 1801 GAAAAATCTACTGGGGATAAATGTAATCTGAAGACTTGTCTCAGAAAACCCCATCTTATATT 1860
Db 1801 GAAAAATCTACTGGGGATAAATGTAATCTGAAGACTTGTCTCAGAAAACCCCATCTTATATT 1860
Qy 1861 CAAAGAGGCAAAAGCTTAAAGGACTAGTCTGCTGGGGTGAATACCAATGATCCT 1920
Db 1861 CAAAGAGGCAAAAGCTTAAAGGACTAGTCTGCTGGGGTGAATACCAATGATCCT 1920
Qy 1921 GAAAAACAGAAAGAAATTTGAAGAACTTTGGAGTTAATGTCTTAATTTATGATAGGATATAT 1980
Db 1921 GAAAAACAGAAAGAAATTTGAAGAACTTTGGAGTTAATGTCTTAATTTATGATAGGATATAT 1980
Qy 1981 GATTTGGATCCCTGAAACCAAAATATATTTCAAAGTGGAGCAATTTGGAACCCCTGAAGCAG 2040
Db 1981 GATTTGGATCCCTGAAACCAAAATATATTTCAAAGTGGAGCAATTTGGAACCCCTGAAGCAG 2040
Qy 2041 GAATTCGCAGAGCTTAAAGAGCTTTTGTGTCCCACTGTTAGCCGCTTTTGTCCCTCATCT 2100
Db 2041 GAATTCGCAGAGCTTAAAGAGCTTTTGTGTCCCACTGTTAGCCGCTTTTGTCCCTCATCT 2100
Qy 2101 TTTGTGGGGAGTCTGATATCCATGTGGATGCAACCGCATTTGATAACGTGGAGAAATGCT 2160
Db 2101 TTTGTGGGGAGTCTGATATCCATGTGGATGCAACCGCATTTGATAACGTGGAGAAATGCT 2160
Qy 2161 TAGTTTTTATTTGCAACAGAGGTCAATTTTGGGGGCTGCACCGCTGTTCTGGGTATTCAATTT 2220
Db 2161 TAGTTTTTATTTGCAACAGAGGTCAATTTTGGGGGCTGCACCGCTGTTCTGGGTATTCAATTT 2220
Qy 2221 TTCATCACTGAGCAATTTGATGCTATGCTTCTGAGTTCCTCAGTTCAATGAAGCAATAA 2280
Db 2221 TTCATCACTGAGCAATTTGATGCTATGCTTCTGAGTTCCTCAGTTCAATGAAGCAATAA 2280
Qy 2281 TGAAGTATTTAACTCTTTTCACTACAGTCTTGTCAAGTATGCTATTTAAATTTACTTTGGCCA 2340
Db 2281 TGAAGTATTTAACTCTTTTCACTACAGTCTTGTCAAGTATGCTATTTAAATTTACTTTGGCCA 2340
Qy 2341 GGTATAAATTTGCCAGTCTCTTTTATAGTGAAAAAATTTATTTGGTTAGTATATAAATA 2400
Db 2341 GGTATAAATTTGCCAGTCTCTTTTATAGTGAAAAAATTTATTTGGTTAGTATATAAATA 2400
Qy 2401 TTTTAAACTTAAATATATAAATCTATATTTTAAACATATGTTCCATTTAAAGCATAGCACT 2460
Db 2401 TTTTAAACTTAAATATATAAATCTATATTTTAAACATATGTTCCATTTAAAGCATAGCACT 2460
Qy 2461 TTGAAATTTAACTATATAAATAGCTCATATTTTACACTTTACAGCTTTTCAATTTGATCAGGTC 2520
Db 2461 TTGAAATTTAACTATATAAATAGCTCATATTTTACACTTTACAGCTTTTCAATTTGATCAGGTC 2520
Qy 2521 TGAAATCTTTTAGCACTTTAAGGAAAAATGACTATGCTAATAATTTATACCTGACCAATGAAAA 2580
Db 2521 TGAAATCTTTTAGCACTTTAAGGAAAAATGACTATGCTAATAATTTATACCTGACCAATGAAAA 2580
Qy 2581 TAAGTACCTCAAAATGATGCAATTTGCACTGGTGAATTTCCAACTGCACAAATCTTTTGTCCT 2640
Db 2581 TAAGTACCTCAAAATGATGCAATTTGCACTGGTGAATTTCCAACTGCACAAATCTTTTGTCCT 2640
Qy 2641 TCTTGTATATAGGTATTTTTTACATGGTTTGAATGCAACACCACTTTTTCATTTTCAATTCAGT 2700
Db 2641 TCTTGTATATAGGTATTTTTTACATGGTTTGAATGCAACACCACTTTTTCATTTTCAATTCAGT 2700
Qy 2701 ATGAACCTTGGAGCTGCTGCCATTTTTTCCACTTTAAACCAACCCAGCTTGAAGGTGAACCTC 2760
Db 2701 ATGAACCTTGGAGCTGCTGCCATTTTTTCCACTTTAAACCAACCCAGCTTGAAGGTGAACCTC 2760
Qy 2761 GAAACTTGTGTTTCAATAATCTTTTCAAAAGTGTGTTTTTACATCAATGTTTAAATTTTCAAAATG 2820
Db 2761 GAAACTTGTGTTTCAATAATCTTTTCAAAAGTGTGTTTTTACATCAATGTTTAAATTTTCAAAATG 2820

QY 971 TCATGACGAGAAATCCGGGAAACAATAGCAGAGTGAGAGTTGACTATATAATTATTA 1030
Db 1047 TCATGACGAGAAATTCGGGAAACAATAGCAGAGTGAGAGTTGACTATATAATTATTA 1106
QY 1031 AGCCATTACAGGATACAGTTGTGACATGAATCTTCAATTTCCAAAGTATGGAAGCCAA 1090
Db 1107 AGCCATTACAGGATACAGTTGTGACATGAATCTTCAATTTCCAAAGTATGGAAGCCAA 1166
QY 1091 GAATACCAATTGGATTTGGCCATCGAGGTGCAGGAACTCTACAAACAATGCCAGCTGG 1150
Db 1167 GAATACCAATTGGATTTGGCCATCGAGGTGCAGGAACTCTACAAACAATGCCAGCTGG 1226
QY 1151 CTAAGTTCAAGAAATACTATTGCTCTTTAAGAAATGCTGTAGTCAATGCTGACCT 1210
Db 1227 CTAAGTTCAAGAAATACTATTGCTCTTTAAGAAATGCTGTAGTCAATGCTGACCT 1286
QY 1211 TTGTAGAAATTTGACGTACACCTTTCAAGGACTTTGTCGCCGTGATATCATCATCTTA 1270
Db 1287 TTGTAGAAATTTGACGTACACCTTTCAAGGACTTTGTCGCCGTGATATCATCATCTTA 1346
QY 1271 CCTGTGTTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTTTGAATTC 1330
Db 1347 CCTGTGTTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTTTGAATTC 1406
QY 1331 CAGTAAAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGCTGAC 1390
Db 1407 CAGTAAAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGCTGAC 1466
QY 1391 TGAATCTTAAGATCGGAAGAACTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAATC 1450
Db 1467 TGAATCTTAAGATCGGAAGAACTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAATC 1526
QY 1451 AGCCATTTCTCTTAAGATGTTTTAGAGTCTTTTGCCAGAAAGATGTAGGGTTAAACA 1510
Db 1527 AGCCATTTCTCTTAAGATGTTTTAGAGTCTTTTGCCAGAAAGATGTAGGGTTAAACA 1586
QY 1511 TTGAATAAATGGAATCTGCCAGCAAGGATGGAATGTGGATGGTAACTTATCAACAT 1570
Db 1587 TTGAATAAATGGAATCTGCCAGCAAGGATGGAATGTGGATGGTAACTTATCAACAT 1646
QY 1571 ATTTTGACATGATCTGTTTTGGATATAATTTTAAACCTGTTTTAGAAATTCGGA 1630
Db 1647 ATTTTGACATGATCTGTTTTGGATATAATTTTAAACCTGTTTTAGAAATTCGGA 1706
QY 1631 AGAGGAGAAATAGTGTCTTCAATTTGATGCAGATATTTGCACAAATGGTTCCGCAAAAGC 1690
Db 1707 AGAGGAGAAATAGTGTCTTCAATTTGATGCAGATATTTGCACAAATGGTTCCGCAAAAGC 1766
QY 1691 AGAACAAATATCCGATACATAATTTTAACTCAAGGAAATCTGAGATTTATCCTGAATCA 1750
Db 1767 AGAACAAATATCCGATACATAATTTTAACTCAAGGAAATCTGAGATTTATCCTGAATCA 1826
QY 1751 TGGACTCAGATCTCGGCAACCCCATCTGCAATGAGCTTTGCACAGTTTGAATCTTAC 1810
Db 1827 TGGACTCAGATCTCGGCAACCCCATCTGCAATGAGCTTTGCACAGTTTGAATCTTAC 1886
QY 1811 TGGGATTAATGACATCTGAAGACTTGTCTCAGAAACCCATCTTATTTCAAGAGGCAA 1870
Db 1887 TGGGATTAATGACATCTGAAGACTTGTCTCAGAAACCCATCTTATTTCAAGAGGCAA 1946
QY 1871 AAGCTAAGGGAATAGTCAATTTCTGCTGGGTGATGATACCAATGATCCTGAAAAACAGAA 1930
Db 1947 AAGCTAAGGGAATAGTCAATTTCTGCTGGGTGATGATACCAATGATCCTGAAAAACAGAA 2006
QY 1931 GGAATTTGAAGAACTTGGAGTTAATGGTCTAATTTTATGATAGGATATATGATTGGATGC 1990
Db 2007 GGAATTTGAAGAACTTGGAGTTAATGGTCTAATTTTATGATAGGATATATGATTGGATGC 2066
QY 1991 CTGAACCAACAAATATATTTCCAAAGTGGAGCAATTCGMAAGCCTGAAGCAGGAATTCGCAG 2050
Db 2067 CTGAACCAACAAATATATTTCCAAAGTGGAGCAATTCGMAAGCCTGAAGCAGGAATTCGCAG 2126
QY 2051 AGCTTAAAGAGCTGTTGTGTGCCACTGTTAGCGGCTTTGTTCCTCATCTTTGTGTGGGG 2110

Db 2127 AGCTTAAAGAGCTGTTGTGTGCCACTGTTAGCGCTTTGTTCCTCATCTTTGTGTGGGG 2186
QY 2111 AGTCTGATATCCATGTGGATGCCAACCGCATTGATAACGTGAGAAATGCTTAGTATTTAT 2170
Db 2187 AGTCTGATATCCATGTGGATGCCAACCGCATTGATAACGTGAGAAATGCTTAGTATTTAT 2246
QY 2171 TGCACAGAGGTCAATTTGGGGCGGTGCACCGCTGTTCTGGGTATTTCATTTTTCACACTG 2230
Db 2247 TGCACAGAGGTCAATTTGGGGCGGTGCACCGCTGTTCTGGGTATTTCATTTTTCACACTG 2306
QY 2231 AGCATTTGTTGATCTATGCTTTTGGCTTTCTCAGTTCAATGAAGCAATAATGAAGTATTT 2290
Db 2307 AGCATTTGTTGATCTATGCTTTTGGCTTTCTCAGTTCAATGAAGCAATAATGAAGTATTT 2366
QY 2291 AACTCTTTCACTACAGTCTTTCGCAAGTATGCTATTAAATTAATTAATTAATTAATTTG 2350
Db 2367 AACTCTTTCACTACAGTCTTTCGCAAGTATGCTATTAAATTAATTAATTAATTAATTTG 2426
QY 2351 CAGTCAAGTCTCTTTATAGTGAGAAATTTATTTGGTTAGTAAATATAAATATTAAACTA 2410
Db 2427 CAGTCAAGTCTCTTTATAGTGAGAAATTTATTTGGTTAGTAAATATAAATATTAAACTA 2486
QY 2411 AATATATAAATCTATTAATGTTAAACATATGTTCAATTAAGCATAGCACTTTGAAATTA 2470
Db 2487 AATATATAAATCTATTAATGTTAAACATATGTTCAATTAAGCATAGCACTTTGAAATTA 2546
QY 2471 CTATATAAATAGCTCATATTTACATTTACATTTTCAAGCTTTTCAATTTGATCAAGTCTG 2530
Db 2547 CTATATAAATAGCTCATATTTACATTTTCAAGCTTTTCAATTTGATCAAGTCTGAAATCTTT 2606
QY 2531 AGCACTTAAGGAAATGACTATGCTAATTTATACCTGACCATGAAATAAATAGTACCTC 2590
Db 2607 AGCACTTAAGGAAATGACTATGCTAATTTATACCTGACCATGAAATAAATAGTACCTC 2666
QY 2591 AAATGATGATGTCATTTGCACTGGTGAATTTCAACTGCACAAATCTTTGTGCCATCTGTATAT 2650
Db 2667 AAATGATGATGTCATTTGCACTGGTGAATTTCAACTGCACAAATCTTTGTGCCATCTGTATAT 2726
QY 2651 AGGTATTTTTTACATGGGTGACATGCACACCAACATTTTCAATTTCAATTTGATGAACCTTG 2710
Db 2727 AGGTATTTTTTACATGGGTGACATGCACACCAACATTTTCAATTTCAATTTGATGAACCTTG 2786
QY 2711 AGGCTGCTGCCATTTTTCACCTTAAACCAACCGCTGAGGTGAACCTCGAAATCTGTT 2770
Db 2787 AGGCTGCTGCCATTTTTCACCTTAAACCAACCGCTGAGGTGAACCTCGAAATCTGTT 2846
QY 2771 TCATAAATCTTTCAAAAGTTGTTTTTACATCAATGTTTAAATTTCAAAATGCTCGAGGTA 2830
Db 2847 TCATAAATCTTTCAAAAGTTGTTTTTACATCAATGTTTAAATTTCAAAATGCTCGAGGTA 2906
QY 2831 ATTTAATGTATAAATAATTTAGTAAAGAAAGATATGTTTGCATATCTTAGTAGAATAGATC 2890
Db 2907 ATTTAATGTATAAATAATTTAGTAAAGAAAGATATGTTTGCATATCTTAGTAGAATAGATC 2966
QY 2891 ACAACATACAAATTCAAATTCAGTGCAATGCTTTAGGTGTTAAGCATGAGATGTTACATGTT 2950
Db 2967 ACAACATACAAATTCAAATTCAGTGCAATGCTTTAGGTGTTAAGCATGAGATGTTACATGTT 3026
QY 2951 TACTGTTAGGTCTTTCGATCTGCTGCTAGGTGAGTATGAGAAAGATGTCAGAGACTGGA 3010
Db 3027 TACTGTTAGGTCTTTCGATCTGCTGCTAGGTGAGTATGAGAAAGATGTCAGAGACTGGA 3086
QY 3011 CGTATTTGTTGCTTAAATAAAGGCTGTTTGTAGCGCTTTTAAATATGCTTATTTTGT 3070
Db 3087 CGTATTTGTTGCTTAAATAAAGGCTGTTTGTAGCGCTTTTAAATATGCTTATTTTGT 3146
QY 3071 TGTGCTCTCACTACCTTATACACATGTTGCTTTGTGGGTTTGTGTTGATGTCGCTGT 3130
Db 3147 TGTGCTCTCACTACCTTATACACATGTTGCTTTGTGGGTTTGTGTTGATGTCGCTGT 3206
QY 3131 GTTATACAGTAGTAAATTTCCATGACAGAAATAAATATGTCCTGAATCTCAAA 3184

Db 3207 GTTATACAGTAGTAAATTTCCATGCAGAAATAAATAATGCTCTGAATTCCTCAT 3260

RESULT 5
AB037855
LOCUS 5443 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1434 protein, partial cds.
ACCESSION AB037855
VERSION
KEYWORDS AB037855.1 GI:7243265
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5443)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Nagase.T., Kikuno.R., Ishikawa.K.I., Hirose.M., and Ohara.O.
XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)

JOURNAL MEDLINE
PUBMED 10718198
REFERENCE 2 (bases 1 to 5443)
Ohara.O., Nagase.T. and Kikuno.R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnaif@kazusa.or.jp, URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913, Fax:+81-438-52-3914)

FEATURES
source
1..5443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hh05877"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
1..5443
/gene="KIAA1434"
-190..2223
/gene="KIAA1434"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1434 protein"
/protein_id="BAA92672.1"
/db_xref="GI:7243266"
/translation="NKSIRWTPSQVAFIRGTLRPGVEPAICGSCDALGNWNPQNAVALPNDTGESMLWKATVILSRGVSVQYRFGYFLEPKTIGGPGQVIVHKWETHLOPRSTPLESEIIIDGGHNGVETLDSGLTCTQIRLRLHYSEKPPVSIITKKLKKSRPRVLTLEGLDEDDDRVPTVLHKMSLSLEISIDNEFKRHSQCEGVLQPDRTWETSOTMEPNLSELPDFREDSHVVOGDALPGHVGTACLSSSTIASGKSAGILTLPMSNRKTIGKRVVDYIIILKPLFQYSDKMSFSKYKPRIPILPDVHRGAGNSTTTLQAKVQENTISLRNASHGAAPVEFDVHLKDFVPVVDHDLTCLTMMKXFDADEVLEFVPEKSLTDFQQLKLTHTVTKSKDKSVQVESENSENPFPSELMVLBSLPDVGFINBIKVIQORDGMDGNLSTYDFMNLFLDIILKTVLENSGKRRIVFSSFDADICTVMROKONKYPIFLFTQKSEIYPELMDLRSTTPAMSAFOENLGINVHTEDLLRNPSYIOEAKAGLIVFCWGDNDNDPENRKLKELGVNGLIYDRIYDMPEQPNFQVEQLERLQELPELXSLCPTVSRVPSLSGSDIHDVNDANGINDVENA"

Query Match 98.8%; Score 3166; DB 9; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3169; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 11 GCGTCGGCGGAGGACGACGCGCGCGCGCGGTACCTCTGCGCGCGGTCTCGCTCTC 70
DB 71 GAGCGCTGGCGGAGGACGACGCGCGCGCGGTACCTCTGCGCGCGGTCTCGCTCTC 130
QY 71 GGGCGGGCGGCGGACGCGGACCTCGGACTAGCGACCCGCGGACGACCATATAA 130
DB 131 GGGCGGGCGGCGGACGCGGACCTCGGACTAGCGACCCGCGGACGACCATATAA 190

131 ATAAATCCATCAGATGACACCTTCTCAGGTTGCCCTTTGAAATTAAGAGAACTCTTTTAC 190
191 ATAAATCCATCAGATGACACCTTCTCAGGTTGCCCTTTGAAATTAAGAGAACTCTTTTAC 250
191 CAGGAGAAAGTTTTCGATATGTGGAAGCTGTGATGCTTTTGGGAACTGGAAATCCTCAAA 250
251 CAGGAGAAAGTTTTCGATATGTGGAAGCTGTGATGCTTTTGGGAACTGGAAATCCTCAAA 310
251 ATGCTGTGGCTCTTCTCCAGAGAAATGACACAGGTTGAAAGCATGCTATGGAAGCAACCA 310
311 ATGCTGTGGCTCTTCTCCAGAGAAATGACACAGGTTGAAAGCATGCTATGGAAGCAACCA 370
311 TTGTTACTCAGTAGAGGAGTATCAGTTTCTGATCGTACTTCAAAGGGTACTTTTTAGAAC 370
371 TTGTTACTCAGTAGAGGAGTATCAGTTTCTGATCGTACTTCAAAGGGTACTTTTTAGAAC 430
371 CAAAGACTATCGGTGGTCCATGTCAGTCAAGTGATAGTTTCAAGTGGGAGACTCATCTAAC 430
431 CAAAGACTATCGGTGGTCCATGTCAGTCAAGTGATAGTTTCAAGTGGGAGACTCATCTAAC 490
431 CACGATCAATAACCCCTTTAGAAAGCGAAATTTATTTTGCAGATGGACAATTTGGAATCC 490
491 CACGATCAATAACCCCTTTAGAAAGCGAAATTTATTTTGCAGATGGACAATTTGGAATCC 550
491 ACAATGGTGTGAAACTCTGATTTCTGGATGGCTGACATGTCAGACTGAAATAAGATTAC 550
551 ACAATGGTGTGAAACTCTGATTTCTGGATGGCTGACATGTCAGACTGAAATAAGATTAC 610
551 GTTTGCATTTCTGAAAGAACTCTCTGTGTCATTAACCAAGAAAGAAATTTAAAGAAATCTA 610
611 GTTTGCATTTCTGAAAGAACTCTCTGTGTCATTAACCAAGAAAGAAATTTAAAGAAATCTA 670
611 GATTTAGGGTGAAGCTGACACTAGAAAGCCCTGGAGAAAGATGACGATGATGGGTATCTC 670
671 GATTTAGGGTGAAGCTGACACTAGAAAGCCCTGGAGAAAGATGACGATGATGGGTATCTC 730
671 CCATGTTACTCCAAAATGTCCTTGGAGATATCCTTTAATAGCGCAATGAGT 730
731 CCATGTTACTCCAAAATGTCCTTGGAGATATCCTTTAATAGCGCAATGAGT 790
731 TCAGTGCAGGCATTCACAGCGGAGTGTGTTAGGCTTCAGCTCATCGTTGGACAG 790
791 TCAGTGCAGGCATTCACAGCGGAGTGTGTTAGGCTTCAGCTCATCGTTGGACAG 850
791 AGTACAGATACAGACGATGGAAACAGATAACCTTGGAACTAATCTTTGATTTTTCGAAG 850
851 AGTACAGATACAGACGATGGAAACAGATAACCTTGGAACTAATCTTTGATTTTTCGAAG 910
851 AAGATCTCAGTGAGCAGTATGTTCAAGGTGATGCCCTTCTGGACATGTGGGTACAGCTT 910
911 AAGATCTCAGTGAGCAGTATGTTCAAGGTGATGCCCTTCTGGACATGTGGGTACAGCTT 970
911 GTCTCTTATCATCCACCATTCCTGAGTGGAAAGAGTCTGGAATTTCTTACTCTTCCCA 970
971 GTCTCTTATCATCCACCATTCCTGAGTGGAAAGAGTCTGGAATTTCTTACTCTTCCCA 1030
971 TCATGAGCAGAAATTTCCCGGAAAAACAATAGCAGAAAGTGGAGTTGACTATATATTA 1030
1031 TCATGAGCAGAAATTTCCCGGAAAAACAATAGCAGAAAGTGGAGTTGACTATATATTA 1090
1031 AGCCATTACAGGATACAGTTGTGACATGAAATCTTCAATTTTCCAAGTATTGGAAGCAAA 1090
1091 AGCCATTACAGGATACAGTTGTGACATGAAATCTTCAATTTTCCAAGTATTGGAAGCAAA 1150
1091 GAATACCATTTGGCTTGGCCATCGAGGTGAGGAACTCTACACAACTGCCAGCTGG 1150
1151 GAATACCATTTGGCTTGGCCATCGAGGTGAGGAACTCTACACAACTGCCAGCTGG 1210
1151 CTAAAGTTCAAGAAATACTATTTGCTTTTAAAGAAATGCTGCTAGTCAATGTCAGCT 1210
1211 CTAAAGTTCAAGAAATACTATTTGCTTTTAAAGAAATGCTGCTAGTCAATGTCAGCT 1270
1211 TTGTAGAAATTTGACGTACACCTTTCAAAGGACTTTTGTGCCCGGTGATATCATGATCTTA 1270

Db	1271	 TTGTAGAAATTTGACGTACACCTTTCAAGAGACTTTGTGCCGTGGTATATCATGATCTTA	1330
Qy	1271	CCTGTTGTTTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTCGAATTAATTTGAAATTC	1330
Db	1331	CCTGTTGTTTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTCGAATTAATTTGAAATTC	1390
Qy	1331	CAGTAAAGAAATTAACAATTTGACCAACTCAGTTGTTAAAGCTCACTCATGTGACTGCAC	1390
Db	1391	CAGTAAAGAAATTAACAATTTGACCAACTCAGTTGTTAAAGCTCACTCATGTGACTGCAC	1450
Qy	1391	TGAATCTAAGGATCGGAAGAACTGTGGTTCAGGAGGAAATTCCTTTTCAGAAAAATC	1450
Db	1451	TGAATCTAAGGATCGGAAGAACTGTGGTTCAGGAGGAAATTCCTTTTCAGAAAAATC	1510
Qy	1451	AGCCATTTCCCTTCTCTTAAGATGGTTTAAAGTCTTTGCGACGAAGATGTAGGGTTTAAACA	1510
Db	1511	AGCCATTTCCCTTCTCTTAAGATGGTTTAAAGTCTTTGCGACGAAGATGTAGGGTTTAAACA	1570
Qy	1511	TTGAAATTAATAATGATCTGCAGCAAGGATCGGAATGTGGATCGTAACTTATCAACAT	1570
Db	1571	TTGAAATTAATAATGATCTGCAGCAAGGATCGGAATGTGGATCGTAACTTATCAACAT	1630
Qy	1571	ATTTTGTGACATGAATCTGTTTGTGGATATAATTTTAAAAAACTGTTTATAGAAAAATCTCGGA	1630
Db	1631	ATTTTGTGACATGAATCTGTTTGTGGATATAATTTTAAAAAACTGTTTATAGAAAAATCTCGGA	1690
Qy	1631	AGAGGAGAATAGTGTTCCTTTCATTTGATCGAGATATTTTGACAAATGTGGTTCGGCAAAAGC	1690
Db	1691	AGAGGAGAATAGTGTTCCTTTCATTTGATCGAGATATTTTGACAAATGTGGTTCGGCAAAAGC	1750
Qy	1691	AGAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGAACTCA	1750
Db	1751	AGAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGAACTCA	1810
Qy	1751	TGGACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTTGACAGATTTGAAAAATCTTAC	1810
Db	1811	TGGACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTTGACAGATTTGAAAAATCTTAC	1870
Qy	1811	TGGGATAAATGTACATCTGAAGACTTGCTCAGAAACCCCATCCTATATTCAGAGGCCAA	1870
Db	1871	TGGGATAAATGTACATCTGAAGACTTGCTCAGAAACCCCATCCTATATTCAGAGGCCAA	1930
Qy	1871	AGCTTAAGGGCTAGTCTATATTCCTGGGGTGAATGATACCAATGATCTCTGAAAAACAGAA	1930
Db	1931	AGCTTAAGGGCTAGTCTATATTCCTGGGGTGAATGATACCAATGATCTCTGAAAAACAGAA	1990
Qy	1931	CGAAAAATTGAAGGAACTTGGAGTTAATGCTCTAAATTTTATGATAGATATATGATGGATGC	1990
Db	1991	CGAAAAATTGAAGGAACTTGGAGTTAATGCTCTAAATTTTATGATAGATATATGATGGATGC	2050
Qy	1991	CTGAAACAAACCAATATATTCAGTGGAGCAATTTGGAACCCCTGAAGCAGGAATTCGCAG	2050
Db	2051	CTGAAACAAACCAATATATTCAGTGGAGCAATTTGGAACCCCTGAAGCAGGAATTCGCAG	2110
Qy	2051	AGCTTAAGAGCTGTTGTGTCCCATCTGTAGCCGCTTTGTTCCTCATCTTTGTGGGG	2110
Db	2111	AGCTTAAGAGCTGTTGTGTCCCATCTGTAGCCGCTTTGTTCCTCATCTTTGTGGGG	2170
Qy	2111	AGTCTGATATCCATGTGATGCCAAACGGCAATGATTAACGTGGAGAAATGCTTAGTTTAT	2170
Db	2171	AGTCTGATATCCATGTGATGCCAAACGGCAATGATTAACGTGGAGAAATGCTTAGTTTAT	2230
Qy	2171	TGCA CAGAGGTCA TTTTGGGGGGGTGCA CCGCTGTTCTGGGTAATTCATTTTTCATCACTG	2230
Db	2231	TGCA CAGAGGTCA TTTTGGGGGGGTGCA CCGCTGTTCTGGGTAATTCATTTTTCATCACTG	2290
Qy	2231	AGCATTTGTTGATCTATGCCCTTTTGGGCTCTCAGTTCGAATGAAGCAATTAATGAAGTATTT	2290
Db	2291	AGCATTTGTTGATCTATGCCCTTTTGGGCTCTCAGTTCGAATGAAGCAATTAATGAAGTATTT	2350
Qy	2291	AACCTCTTTCATCAGTCTTCGCAAGTATGCTATTTAAATTTACTTTGGCCAGGTATATTCG	2350

Db	2351	AACTCTTCACTACAGTCTTGGCAAGTATGCTATTTAAATTAATCTGGCCAGGTATPAATTG	2411
Qy	2351	CCAGTCAGTCTCTTTATAGTGAGAAAAATTTATTTGGTTAGTAAATATAAAATATTTTAAACTA	2410
Db	2411	CCAGTCAGTCTCTTTATAGTGAGAAAAATTTATTTGGTTAGTAAATATAAAATATTTTAAACTA	2470
Qy	2411	AAATATAAAATCTATAAATGTTAAACATATGTTTCATTTAAAGCATAGCACCTTTGAAATTA	2470
Db	2471	AAATATAAAATCTATAAATGTTAAACATATGTTTCATTTAAAGCATAGCACCTTTGAAATTA	2530
Qy	2471	CTATATAAAATAGCTCATATTTTACACATTTTACAGCTTTTTCATTTGATCAGGTCGTAATCTTTT	2530
Db	2531	CTATATAAAATAGCTCATATTTTACACATTTTACAGCTTTTTCATTTGATCAGGTCGTAATCTTTT	2590
Qy	2531	AGCACTTTAAGGAAAAATGACTATGCATAATATATACCTGACCAATGAAAAAAATAAGTACCTC	2590
Db	2591	AGCACTTTAAGGAAAAATGACTATGCATAATATATACCTGACCAATGAAAAAAATAAGTACCTC	2650
Qy	2591	AAATGCAATGATTTGCACTGGTGTATCCCACTGCACAAAATCTTTGTGCCATCTTGATATAT	2650
Db	2651	AAATGCAATGATTTGCACTGGTGTATCCCACTGCACAAAATCTTTGTGCCATCTTGATATAT	2710
Qy	2651	AGGTATTTTTTACATGCTGTTGACATGCACACCAACCATTTTTCATTCAGTATGAACCTTG	2710
Db	2711	AGGTATTTTTTACATGCTGTTGACATGCACACCAACCATTTTTCATTCAGTATGAACCTTG	2770
Qy	2711	AGGCTGCTGCCATTTTCCACTTAACCAAAACCGACCTGAAAGGTGAACCTCGAAACTTGTT	2770
Db	2771	AGGCTGCTGCCATTTTCCACTTAACCAAAACCGACCTGAAAGGTGAACCTCGAAACTTGTT	2830
Qy	2771	TCATATAATCTTTCAAAAAGTTGTTTATCATCAATGTTTAAAAATTTTCAAAAATGCTGCAGGTA	2830
Db	2831	TCATATAATCTTTCAAAAAGTTGTTTATCATCAATGTTTAAAAATTTTCAAAAATGCTGCAGGTA	2890
Qy	2831	ATTTAAATCTATAAAATATTTAGTAAAGAAAAAGTATGTTATGTCATCTTTAGTAGAATAGATC	2890
Db	2891	ATTTAAATCTATAAAATATTTAGTAAAGAAAAAGTATGTTATGTCATCTTTAGTAGAATAGATC	2950
Qy	2891	ACAACATACAAATCAATTCAGTGCATGCTTTTAGGTGTTAAGCATGAGATGTACATGTT	2950
Db	2951	ACAACATACAAATCAATTCAGTGCATGCTTTTAGGTGTTAAGCATGAGATGTACATGTT	3010
Qy	2951	TACTGTTAGGTCCTGTCATCTGCTGCTAGTGTAGTATGAGAAGATGCTCAAGGACTGGA	3010
Db	3011	TACTGTTAGGTCCTGTCATCTGCTGCTAGTGTAGTATGAGAAGATGCTCAAGGACTGGA	3070
Qy	3011	CGTATTTTGTTCCTTAAAAAAGGCTGTTGTAGCGCTTTTAAATATGCTTATTTTG	3070
Db	3071	CGTATTTTGTTCCTTAAAAAAGGCTGTTGTAGCGCTTTTAAATATGCTTATTTTG	3130
Qy	3071	TGTGTCCTCTACTACTATTTACACATGCTGCTTTGCGGTTTGTGTTGATGTCGCTGT	3130
Db	3131	TGTGTCCTCTACTACTATTTACACATGCTGCTTTGCGGTTTGTGTTGATGTCGCTGT	3190
Qy	3131	GTATACAGTAGTAAATTTTCCATGCAGAAAAATAAATGCTCTGAAATCTCAAA	3184
Db	3191	GTATACAGTAGTAAATTTTCCATGCAGAAAAATAAATGCTCTGAAATCTCAAA	3244
RESULT 6			
CQ491863			
LOCUS	CQ491863 3499 bp DNA		
DEFINITION	Sequence 23730 from Patent WO0160860.		
ACCESSION	CQ491863		
VERSION	CQ491863.1		
KEYWORDS	GI:41457482		
SOURCE	Homo sapiens (human)		
	linear	PAT 30-JAN-2000	

RESULT	6
LOCUS	CQ491863
DEFINITION	Sequence 23730 from Patent WO0160860.
ACCESSION	CQ491863
VERSION	CQ491863.1 GI:41457482
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1
REFERENCE	Schlegel,R., Endege,W.O. and Monahan,J.E.
AUTHORS	Genes differentially expressed in human prostate cancer and their TITLE
PATENT	PAT 30-JAN-2000 linear

Qy 1991 CTGAACCAACCAATATATTCCTAGTGGAGCAATTTGGAAACGCTTGAAGCAGGAATTCGCCAG 2050
Db 2067 CTGAACCAACCAATATATTCCTAGTGGAGCAATTTGGAAACGCTTGAAGCAGGAATTCGCCAG 2126
Qy 2051 AGCTTAAGAGCTGTTGTGTGCCACTGTAGCCGCTTTGTTCCTCATCTTTCTGTGGGG 2110
Db 2127 AGCTTAAGAGCTGTTGTGTGCCACTGTAGCCGCTTTGTTCCTCATCTTTCTGTGGGG 2186
Qy 2111 AGTCTGATATCCATGTGGATGCCAAACGGCATTTGATAACGTTGGAGATGCTTAGTTTTAT 2170
Db 2187 AGTCTGATATCCATGTGGATGCCAAACGGCATTTGATAACGTTGGAGATGCTTAGTTTTAT 2246
Qy 2171 TGCACAGAGTCATTTTGGGGGCTGCACCGCTGCTTGTGGGTATTCATTTTCACTCACTG 2230
Db 2247 TGCACAGAGTCATTTTGGGGGCTGCACCGCTGCTTGTGGGTATTCATTTTCACTCACTG 2306
Qy 2231 AGCATTTGTATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAATAATGAAGTATTT 2290
Db 2307 AGCATTTGTATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAATAATGAAGTATTT 2366
Qy 2291 AACTCTTTTCACTACAGTTCTTTCGAAGTATGCTATTTAAATTACTTGCCGAGGTATTAATG 2350
Db 2367 AACTCTTTTCACTACAGTTCTTTCGAAGTATGCTATTTAAATTACTTGCCGAGGTATTAATG 2426
Qy 2351 CCAGTCAGTCTCTTTATAGTGAGAAATTTATTTGGTTAGTAAATATAATATTTTAAACTA 2410
Db 2427 CCAGTCAGTCTCTTTATAGTGAGAAATTTATTTGGTTAGTAAATATAATATTTTAAACTA 2486
Qy 2411 AATATATAATCTATAATGTTTAAACATATGTTTCAATTAAGCAATAGACATTTTGAATTA 2470
Db 2487 AATATATAATCTATAATGTTTAAACATATGTTTCAATTAAGCAATAGACATTTTGAATTA 2546
Qy 2471 CTATATAATAGTCTCATATTTTACCTTACAGCTTTTCAATTTTGAATCAAGTCTGAAATCTTT 2530
Db 2547 CTATATAATAGTCTCATATTTTACCTTACAGCTTTTCAATTTTGAATCAAGTCTGAAATCTTT 2606
Qy 2531 AGCATTAAAGAAATAGCATGATTAATTTACCTGACCAATGAAGCAATTAAGTACCTC 2590
Db 2607 AGCATTAAAGAAATAGCATGATTAATTTACCTGACCAATGAAGCAATTAAGTACCTC 2666
Qy 2591 AAATGATGCAATTTGACCTGTTGATTTCCAACTGACCAATCTTTTGTGCACTTTGTATAT 2650
Db 2667 AAATGATGCAATTTGACCTGTTGATTTCCAACTGACCAATCTTTTGTGCACTTTGTATAT 2726
Qy 2651 AGTATATTTTACATGGGTTGATGCAACCAACCAATTTTCAATTCAGTATGAACCTTG 2710
Db 2727 AGTATATTTTACATGGGTTGATGCAACCAACCAATTTTCAATTCAGTATGAACCTTG 2786
Qy 2711 AGGCTGCTGCCATTTTCCACTTAAACCAACAGGCTGAGGTGAACCTCGAATCTGTT 2770
Db 2787 AGGCTGCTGCCATTTTCCACTTAAACCAACAGGCTGAGGTGAACCTCGAATCTGTT 2846
Qy 2771 TCATAAATCTTTCAAAAGTTGTTTATACATCAATGTTTAAATTTCAAAATGCTGACGGTA 2830
Db 2847 TCATAAATCTTTCAAAAGTTGTTTATACATCAATGTTTAAATTTCAAAATGCTGACGGTA 2906
Qy 2831 ATTAAATGATATAAATATTAGTAAAGAAAGTATGTTTGCATACTTAGTAAATAGATC 2890
Db 2907 ATTAAATGATATAAATATTAGTAAAGAAAGTATGTTTGCATACTTAGTAAATAGATC 2966
Qy 2891 ACAACATACAAATTTCAATTCAGTGAATGTTTAGGTGTTAAGCATGAGTATGATCATGTT 2950
Db 2967 ACAACATACAAATTTCAATTCAGTGAATGTTTAGGTGTTAAGCATGAGTATGATCATGTT 3026
Qy 2951 TACTGTTAGTCTCTGCACTGCTGCTAGGTGAGTATGAGAGATGTCAGGACTGGA 3010
Db 3027 TACTGTTAGTCTCTGCACTGCTGCTAGGTGAGTATGAGAGATGTCAGGACTGGA 3086
Qy 3011 CGTATTTTGTGCTAAAAAAGGCTGTTTGTAGGGCTTTTAAATATGCTATTTTGT 3070
Db 3087 CGTATTTTGTGCTAAAAAAGGCTGTTTGTAGGGCTTTTAAATATGCTATTTTGT 3146
Qy 3071 TGTGCTCTCACTACCTATTTACACATGTTGCTTTTGTGGGTTTGTGTTGATGTTGCTG 3130

Db 3147 TGTGCTCTCACTACCTATTTACACACTGTTGCTTTGGGTTTGTGTTGATGCGGTGT 3206
Qy 3131 GTTATACAGTATGTTAAATTTCCATGCGAGAAAAATAATGTCTCTGAAATCTCAAA 3184
Db 3207 GTTATACAGTATGTTAAATTTCCATGCGAGAAAAATAATGTCTCTGAAATTTTATA 3260
RESULT 7
CQ497765 3499 bp DNA linear PAT 30-JAN-2004
LOCUS Sequence 29632 from Patent WO0160860.
DEFINITION CQ497765
ACCESSION CQ497765
VERSION CQ497765.1 GI:41463401
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1
REFERENCE
AUTHORS
TITLE
JOURNAL
Schlegel, R., Endege, W.O. and Monahan, J.E.
Genes differentially expressed in human prostate cancer and their
use
Patent: WO 0160860-A 29632 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
LOCATION/Qualifiers
1. 3499
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
FEATURES
source
ORIGIN
Query Match 98.7%; Score 3164.4; DB 6; Length 3499;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3168; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 11 GCCTCCGGCGGACGACGCGCGCGGACCTCTCTCGCGGCTCTCTCGCGGCTCTC 70
Db 87 GAGCGCTGGCGAGGACGCGACGCGCGCGGACCTCTCTCGCGGCTCTCTCGCGCTC 146
Qy 71 GGGCGGGCGGCGGCGACGCGGACCTCTCGGATAGCGGACCGGAGCAGCATATAAA 130
Db 147 GGGCGGGCGGCGGCGACGCGGACCTCTCGGATAGCGGACCGGAGCAGCATATAAA 206
Qy 131 ATAAATCCATCAAAATGACACCTTCTCAGTTCCTTTGAAATAGAGGAACCTTTTAC 190
Db 207 ATAAATCCATCAAAATGACACCTTCTCAGTTCCTTTGAAATAGAGGAACCTTTTAC 266
Qy 191 CAGGAGAAAGTTTTCGATATGTGGAAGCTGTGATGCTTTGGGAACTCGAATCTCTCAA 250
Db 267 CAGGAGAAAGTTTTCGATATGTGGAAGCTGTGATGCTTTGGGAACTCGAATCTCTCAA 326
Qy 251 ATGCTGTGGCTCTCTTTCGAGAAATGACACAGGTGAAGCATGCTATGGAAGCAACCA 310
Db 327 ATGCTGTGGCTCTCTTTCGAGAAATGACACAGGTGAAGCATGCTATGGAAGCAACCA 386
Qy 311 TTGTTACTCAGTACGAGGATGATCAGTTTCTCAGTATCGCTTCAAGGGTACTTTTGAAC 370
Db 387 TTGTTACTCAGTACGAGGATGATCAGTTTCTCAGTATCGCTTCAAGGGTACTTTTGAAC 446
Qy 371 CAAAGACTATCGGTGGTCCATGTCAAGTGA TAGTTTCAAGTGGGAGATCTCTCTCAAC 430
Db 447 CAAAGACTATCGGTGGTCCATGTCAAGTGA TAGTTTCAAGTGGGAGATCTCTCTCAAC 506
Qy 431 CAGCATCAATAACCCCTTTTGAAGCGAAATTTATTTAGCATGGAATTTGGATCC 490
Db 507 CAGCATCAATAACCCCTTTTGAAGCGAAATTTATTTAGCATGGAATTTGGATCC 566
Qy 491 ACAATGTTGTTGAACCTCTGATTTCTGATGGCTGACATGTCAGACTGAAATAGATTAC 550
Db 567 ACAATGTTGTTGAACCTCTGATTTCTGATGGCTGACATGTCAGACTGAAATAGATTAC 626
Qy 551 GTTTGCTCTCACTACCTATTTACACATGTTGCTTTTGTGGGTTTGTGTTGATGTTGCTG 610

||||| 627 GTTTGCAATTAATCTGAAACCTCCCTGTCAATACCAAGAAAAATTAATAAATCTTA 686
QY
||||| 611 GATTTAGGGTGAAGCTGACACCTAGAGGCTCGAGGAAGATGACGATGATGGGTATCTC 670
Db
||||| 687 GATTTAGGGTGAAGCTGACACCTAGAGGCTCGAGGAAGATGACGATGATGGGTATCTC 746
QY
||||| 671 CCACTGTACTCCACAAAATGTCCAATAGCTTGGAGATATCCTTAATAAGCGACAATAGT 730
Db
||||| 747 CCACTGTACTCCACAAAATGTCCAATAGCTTGGAGATATCCTTAATAAGCGACAATAGT 806
QY
||||| 731 TCAAGTGCAGGCATTCACAGCCGAGTGTGGTTATGGCTTGCAGCTGATCGTGCAG 790
Db
||||| 807 TCAAGTGCAGGCATTCACAGCCGAGTGTGGTTATGGCTTGCAGCTGATCGTGCAG 866
QY
||||| 791 AGTACAGCATACAGAGATGGAACACAGATAAACCCTGGAATAATCTTTGATTTTTCGAAG 850
Db
||||| 867 AGTACAGCATACAGAGATGGAACACAGATAAACCCTGGAATAATCTTTGATTTTTCGAAG 926
QY
||||| 851 AAGATCTCAGTGAGCAGTAGTTTCAAGGTGATGCCCTTCTCGACATGFGGTACAGCTT 910
Db
||||| 927 AAGATCTCAGTGAGCAGTAGTTTCAAGGTGATGCCCTTCTCGACATGFGGTACAGCTT 986
QY
||||| 911 GTCTCTTATCATCCACCATTTGCTGAGAGTGGAAAGAGTCTGGAATCTTACTCTCCCA 970
Db
||||| 987 GTCTCTTATCATCCACCATTTGCTGAGAGTGGAAAGAGTCTGGAATCTTACTCTCCCA 1046
QY
||||| 971 TCATGAGCAGAAATTTCCCGGAAAAAATAAGGCAAGTGAAGTGAATATAAATATTA 1030
Db
||||| 1047 TCATGAGCAGAAATTTCCCGGAAAAAATAAGGCAAGTGAAGTGAATATAAATATTA 1106
QY
||||| 1031 AGCCATTTACCAGATACAGTTGTGACATGAATCTTTCAATTTTCCAAGTATGGAAGCCAA 1090
Db
||||| 1107 AGCCATTTACCAGATACAGTTGTGACATGAATCTTTCAATTTTCCAAGTATGGAAGCCAA 1166
QY
||||| 1091 GAATACCATTTGATGTGGCCATCGAGTGCAGGAAACTCTACACAACTGCCAGCTGG 1150
Db
||||| 1167 GAATACCATTTGATGTGGCCATCGAGTGCAGGAAACTCTACACAACTGCCAGCTGG 1226
QY
||||| 1151 CTAAAGTTCAAGAAATACTATTGTCTTTTAAAGAAATGCTGTAAGTCAATGTGCAGCT 1210
Db
||||| 1227 CTAAAGTTCAAGAAATACTATTGTCTTTTAAAGAAATGCTGTAAGTCAATGTGCAGCT 1286
QY
||||| 1211 TTGTAGAAATTTGACGTGACCTTTCAAGGACTTTTGTGCCGTGGTATATCATGATCTTA 1270
Db
||||| 1287 TTGTAGAAATTTGACGTGACCTTTCAAGGACTTTTGTGCCGTGGTATATCATGATCTTA 1346
QY
||||| 1271 CTTGTTGTTGACTATGAAAAAATAATTTGATGCTGATCCAGTTGAAATTTTGAATTC 1330
Db
||||| 1347 CTTGTTGTTGACTATGAAAAAATAATTTGATGCTGATCCAGTTGAAATTTTGAATTC 1406
QY
||||| 1331 CAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCAC 1390
Db
||||| 1407 CAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCAC 1466
QY
||||| 1391 TGAATCTAAGGATCGGAAGAAATCTGTGGTTCAGGAGAAATTCCTTTTCAGAAAAATC 1450
Db
||||| 1467 TGAATCTAAGGATCGGAAGAAATCTGTGGTTCAGGAGAAATTCCTTTTCAGAAAAATC 1526
QY
||||| 1451 AGCCATTTCTCTTAAGATGGTTTATAGATCTTTGCGCAGAAATGTAGGGTTTAACA 1510
Db
||||| 1527 AGCCATTTCTCTTAAGATGGTTTATAGATCTTTGCGCAGAAATGTAGGGTTTAACA 1586
QY
||||| 1511 TTGAAATAAATGGAATCTGCCAGCAAGGATGGAAATGGGATGGTAACTTATCAACAT 1570
Db
||||| 1587 TTGAAATAAATGGAATCTGCCAGCAAGGATGGAAATGGGATGGTAACTTATCAACAT 1646
QY
||||| 1571 ATTTTGACATGAATCTGTTTTTGGATATAATTTTAAAAAATCTGTTTTAGAAAAATCTGGGA 1630
Db
||||| 1647 ATTTTGACATGAATCTGTTTTTGGATATAATTTTAAAAAATCTGTTTTAGAAAAATCTGGGA 1706
QY
||||| 1631 AGAGGAGAAATAGTGTCTTCAATTTGATGACAGATATTTGCAATGTTTCGGCAAAAGC 1690

Db
||||| 1707 AGAGGAGAAATAGTGTCTTCAATTTGATGACAGATATTTGCACAAATGGTTTCGGCAAAAGC 1766
QY
||||| 1691 AGAACAAATATCCGATACTATTTTAACTCAAGGAAAAATCTGAGATTTTATCTGAACTCA 1750
Db
||||| 1767 AGAACAAATATCCGATACTATTTTAACTCAAGGAAAAATCTGAGATTTTATCTGAACTCA 1826
QY
||||| 1751 TGGACCTCAGATCTCGGACAAACCCCATTTGAATGAGCTTTGCACAGTTTGAATATCTAC 1810
Db
||||| 1827 TGGACCTCAGATCTCGGACAAACCCCATTTGAATGAGCTTTGCACAGTTTGAATATCTAC 1886
QY
||||| 1811 TGGGATAAATGTACATCTGAAGCTTTGCTCAGAAAAATCTTATATCTCAAGAGGCA 1870
Db
||||| 1887 TGGGATAAATGTACATCTGAAGCTTTGCTCAGAAAAATCTTATATCTCAAGAGGCA 1946
QY
||||| 1871 AAGCTAAGGGACTAGTCTATATTTCTGCTGGGGTGAATGATACCAATGATCCTGAAAAACAG 1930
Db
||||| 1947 AAGCTAAGGGACTAGTCTATATTTCTGCTGGGGTGAATGATACCAATGATCCTGAAAAACAG 2006
QY
||||| 1931 GGAATTTGAAGAACTTTGGAGTTAATGCTTAATTTATGATAGATATATGATTTGGATGC 1990
Db
||||| 2007 GGAATTTGAAGAACTTTGGAGTTAATGCTTAATTTATGATAGATATATGATTTGGATGC 2066
QY
||||| 1991 CTGAACACCAAAATATATTTCCAAGTGGAGCAATTTGGAACGCTGAAGCAGGAATTTGCCAG 2050
Db
||||| 2067 CTGAACACCAAAATATATTTCCAAGTGGAGCAATTTGGAACGCTGAAGCAGGAATTTGCCAG 2126
QY
||||| 2051 AGCTTAAAGAGCTGTTTGTGCTCCACTGTTAGCCGCTTTGTTCCCTCATCTTTTGTGGGG 2110
Db
||||| 2127 AGCTTAAAGAGCTGTTTGTGCTCCACTGTTAGCCGCTTTGTTCCCTCATCTTTTGTGGGG 2186
QY
||||| 2111 AGTCTGATATCCATGTGGATGCCAAACGGCATGTATTAAGTGGAGAAATGCTTGTATTTAT 2170
Db
||||| 2187 AGTCTGATATCCATGTGGATGCCAAACGGCATGTATTAAGTGGAGAAATGCTTGTATTTAT 2246
QY
||||| 2171 TGCAAGAGGTCATTTTGGGGCGTGACCGCTGCTTGGGTATTTTCAATTTTCACTCAGT 2230
Db
||||| 2247 TGCAAGAGGTCATTTTGGGGCGTGACCGCTGCTTGGGTATTTTCAATTTTCACTCAGT 2306
QY
||||| 2231 AGCATTTGTGATCTATGCTTTTGGGCTTTCTCAGTTCAATGAAGCAATTAATGAATATTT 2290
Db
||||| 2307 AGCATTTGTGATCTATGCTTTTGGGCTTTCTCAGTTCAATGAAGCAATTAATGAATATTT 2366
QY
||||| 2291 AACTCTTTCACTACAGTTCTTGGCAAGTATGCTATTTAAATTAATTTGGCCAGGTATAATG 2350
Db
||||| 2367 AACTCTTTCACTACAGTTCTTGGCAAGTATGCTATTTAAATTAATTTGGCCAGGTATAATG 2426
QY
||||| 2351 CCAGTCAGTCTCTTTATAGTCAGAAAAATTTATGTTAGTAAATATAAATATTTTAAACTA 2410
Db
||||| 2427 CCAGTCAGTCTCTTTATAGTCAGAAAAATTTATGTTAGTAAATATAAATATTTTAAACTA 2486
QY
||||| 2411 AATATAAATCTATAATTTTAAACATATGTTTCAATTTAAAGCATAGCACTTTGAAATTA 2470
Db
||||| 2487 AATATAAATCTATAATTTTAAACATATGTTTCAATTTAAAGCATAGCACTTTGAAATTA 2546
QY
||||| 2471 CTATATAAATAGCTCATATTTTACCTACAGCTTTTCAATTTGATCAGTCTGAAATCTTT 2530
Db
||||| 2547 CTATATAAATAGCTCATATTTTACCTACAGCTTTTCAATTTGATCAGTCTGAAATCTTT 2606
QY
||||| 2531 AGCATTTAAGGAAATGATGATGATTAATTAATGCTGACCAATGAAGAAATTAATGATCCTC 2590
Db
||||| 2607 AGCATTTAAGGAAATGATGATGATTAATTAATGCTGACCAATGAAGAAATTAATGATCCTC 2666
QY
||||| 2591 AAATGCAATGCTTTGCACTGGTGAATTTCAACTGCAAAATCTTTTGTGCACTCTGTATAT 2650
Db
||||| 2667 AAATGCAATGCTTTGCACTGGTGAATTTCAACTGCAAAATCTTTTGTGCACTCTGTATAT 2726
QY
||||| 2651 AGGATTTTATGAGGTGCAATGCAACAAACCAATTTTCAATTTTCAATGATGAACCTTG 2710
Db
||||| 2727 AGGATTTTATGAGGTGCAATGCAACAAACCAATTTTCAATTTTCAATGATGAACCTTG 2786
QY
||||| 2711 AGGCTGCTGCAATTTTCCACTTAACCAACAGCCTGAAGTGAACCTTGAACCTTGT 2770
Db
||||| 2787 AGGCTGCTGCAATTTTCCACTTAACCAACAGCCTGAAGTGAACCTTGAACCTTGT 2846

QY	2771	TCATTAATCTTTCAAAGTTGTTTTTATCATCAATGTTTAAAAATTTCAAAATGCTCGAGGTA	2830
Db	2847	TCATAAATCTTTCAAAGTTGTTTTTATCATCAATGTTTAAAAATTTCAAAATGCTCGAGGTA	2906
QY	2831	ATTTAATGTTATAAAATATTAGTAAGAAAAAGTATGTATTCATACCTACTTAGTAGAATAGATC	2890
Db	2907	ATTTAATGTTATAAAATATTAGTAAGAAAAAGTATGTATTCATACCTACTTAGTAGAATAGATC	2966
QY	2891	ACAACATACAAAATTCAAATCAGTGCGATGCTTTTAGTGCTTTAAGCATGAGATTGTACATGTT	2950
Db	2967	ACAACATACAAAATTCAAATCAGTGCGATGCTTTTAGTGCTTTAAGCATGAGATTGTACATGTT	3026
QY	2951	TACTGTTAGGTCCTTGCATCTGTGGTGCTAGGTGAGTATCGAAGAAGTGTCAAGAGACTGGA	3010
Db	3027	TACTGTTAGGTCCTTGCATCTGTGGTGCTAGGTGAGTATCGAAGAAGTGTCAAGAGACTGGA	3086
QY	3011	CGTATTTGTTGGCCCTAAAAAAAAGGCTGTGTTGTAGGCGGTTTTTAAATATGCTTATTTTG	3070
Db	3087	CGTATTTGTTGGCCCTAAAAAAAAGGCTGTGTTGTAGGCGGTTTTTAAATATGCTTATTTTG	3146
QY	3071	TGTCGTCCTCACTACCTATTACACACATGTTGCTTTTGTGGGTTTGTGTTGATGCGGTGT	3130
Db	3147	TGTCGTCCTCACTACCTATTACACACATGTTGCTTTTGTGGGTTTGTGTTGATGCGGTGT	3206
QY	3131	GTTTATACAGTAGTTAAATTTCCATGCGAGAAAAATAAAATGCTCTCAATTCCTCAA	3184
Db	3207	GTTTATACAGTAGTTAAATTTCCATGCGAGAAAAATAAAATGCTCTCAATTCCTCAA	3260
RESULT 8			
HSW804380		6671 bp	linear
LOCUS	HSW804380	6671 bp	linear
DEFINITION	Homo sapiens mRNA; cDNA DKFZp45101715 (from clone DKFZp45101715).		
ACCESSION	AL833069		
VERSION	AL833069.1	GI:21733660	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Bloeker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY		
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp45101715) is available at the RZPD in Berlin.. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.		
FEATURES	Location/Qualifiers		
source	1..6671		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="RZPD:DKFZp45101715"		
	/db_xref="taxon:9606"		
	/clone="DKFZp45101715"		
	/tissue_type="human spinal cord"		
	/clone_lib="451 (synonym: hiccl). Vector pSport1; host DH10B; sites NotI + SalI"		
	/dev_stage="adult"		
	6630..6635		
	polya_signal		
	polya_site		
ORIGIN	6645		
Query Match	85.1%	Score 2726.8;	DB 9; Length 6671;

Db	2732	TAAGATGGTTTTAGAGTCTTTGGCAGAAAGATGTAGGGGTTTAAACATTTGAAATATAAAATGGAT	2791
Qy	1527	CTGCCAGCAAAAGGATGGAATGTGGGATGGTAACTTATCAACATATATTTTGACATGAATCT	1586
Db	2792	CTGCCAGCAAAAGGATGGAATGTGGGATGGTAACTTATCAACATATATTTTGACATGAATCT	2851
Qy	1587	GTTTTTGGATATAATATTTTAAAAACCTGTTTTAGAAAAATCTGGGGAAGAGAGAAATAGTGTT	1646
Db	2852	GTTTTTGGATATAATATTTTAAAAAACCTGTTTTAGAAAAATCTGGGGAAGAGAGAAATAGTGTT	2911
Qy	1647	TTCTTTCATTTTGATGCAGATATTTGCACAATGGTTTCGGCAAAAGCAGAAACAAATATCCGAT	1706
Db	2912	TTCTTTCATTTTGATGCAGATATTTGCACAATGGTTTCGGCAAAAGCAGAAACAAATATCCGAT	2971
Qy	1707	ACTATTTTTAACTCAAGGAAAATCTGAGATTTATCTCGAACTCATGTGACCTCAGATCTCG	1766
Db	2972	ACTATTTTTAACTCAAGGAAAATCTGAGATTTATCTCGAACTCATGTGACCTCAGATCTCG	3031
Qy	1767	GACAACCCCAATTGCAATGAGCTTTGACAGTTTGAAAAATCTACTGGGGATAAATGTATCA	1826
Db	3032	GACAACCCCAATTGCAATGAGCTTTGACAGTTTGAAAAATCTACTGGGGATAAATGTATCA	3091
Qy	1827	TACTGAGAGCTTGCTCAGNAACCCATCCTATATTTCAAGAGGCAAAAGCTTAAGGACTAGT	1886
Db	3092	TACTGAGAGCTTGCTCAGNAACCCATCCTATATTTCAAGAGGCAAAAGCTTAAGGACTAGT	3151
Qy	1887	CATATTTCTGTGGGCTGATGATACCAATGATCTCTGAAAAACAGAAAGGAAATTTGAAGGAAT	1946
Db	3152	CATATTTCTGTGGGCTGATGATACCAATGATCTCTGAAAAACAGAAAGGAAATTTGAAGGAAT	3211
Qy	1947	TGGAGTTAAATGGTCTAAATTTATGATAGGATATATGATTTGGATGCTTGAAACCAACAAATAT	2006
Db	3212	TGGAGTTAAATGGTCTAAATTTATGATAGGATATATGATTTGGATGCTTGAAACCAACAAATAT	3271
Qy	2007	ATTCCAAGTGGAGCAATTTGGNAACCCCTGAGCAGGAATTTGCCAGAGCTTTAAGAGCTGTTT	2066
Db	3272	ATTCCAAGTGGAGCAATTTGGNAACCCCTGAGCAGGAATTTGCCAGAGCTTTAAGAGCTGTTT	3331
Qy	2067	GTGTCCACACTGTTAGCGCTTTGTTCCTCATCTTTGTGGGAGTCTCATATCCATGT	2126
Db	3332	GTGTCCACACTGTTAGCGCTTTGTTCCTCATCTTTGTGGGAGTCTCATATCCATGT	3391
Qy	2127	GGATGCCAACGGCAATTGATAACGTGGAGAAATGCTTAGTTTTTATTTGCACAGAGTCAATTT	2186
Db	3392	GGATGCCAACGGCAATTGATAACGTGGAGAAATGCTTAGTTTTTATTTGCACAGAGTCAATTT	3451
Qy	2187	TGGGGGGTGACCGCTGTTCTGGGTATTCATTTTTTTCATCACTGAGCAATTTGTGATCTAT	2246
Db	3452	TGGGGGGTGACCGCTGTTCTGGGTATTCATTTTTTTCATCACTGAGCAATTTGTGATCTAT	3511
Qy	2247	GCCTTTTGGGCTTCTCAGTTTCAATGAAGCAATATGAAGTATTTAACTCTTTTCACTACAG	2306
Db	3512	GCCTTTTGGGCTTCTCAGTTTCAATGAAGCAATATGAAGTATTTAACTCTTTTCACTACAG	3571
Qy	2307	TTCTTGAAGTATGCTATTTTAAATTACTTTGGCCAGGTATAATTTGCCAGTCACTCTTTTA	2366
Db	3572	TTCTTGAAGTATGCTATTTTAAATTACTTTGGCCAGGTATAATTTGCCAGTCACTCTTTTA	3631
Qy	2367	TAGTGAGAAAAATTTATTTGGTTAGTAATATAAATATTTTAACTAAATATATAAATCTATA	2426
Db	3632	TAGTGAGAAAAATTTATTTGGTTAGTAATATAAATATTTTAACTAAATATAAATCTATA	3691
Qy	2427	ATGTTAAACATATGTTTCATTTAAAGCATAGCACCTTGAAATTAACCTATAAATAGTCTCA	2486
Db	3692	ATGTTAAACATATGTTTCATTTAAAGCATAGCACCTTGAAATTAACCTATAAATAGTCTCA	3751
Qy	2487	TATTTACACTTACAGCTTTTTCATTTTGATCAGGTCTGAAATCTTTTAGCACCTTAAGGAAAT	2546
Db	3752	TATTTACACTTACAGCTTTTTCATTTTGATCAGGTCTGAAATCTTTTAGCACCTTAAGGAAAT	3811
Qy	2547	GACTATGCATAATTTATACCTGACCATGAAAAAATAAGTACCTCAAAATGCATTCATTTGC	2606
Db	3812	GACTATGCATAATTTATACCTGACCATGAAAAAATAAGTACCTCAAAATGCATTCATTTGC	3871

QY	2607	ACTGGTGATTCACAACTCGCAAAATCTTTGTGCGCAATCTTTGTATATAGGTATTTTACATG	2666
DB	3872	ACTGGTGATTCACAACTCGCAAAATCTTTGTGCGCAATCTTTGTATATAGGTATTTTACATG <td>3931</td>	3931
QY	2667	GGTTGACATGCACACACACCAATTTTCATTCAGTATGAACTTGAGGCTGCTGCATTTT	2726
DB	3932	GGTTGACATGCACACACCAATTTTCATTCAGTATGAACTTGAGGCTGCTGCATTTT	3991
QY	2727	TCCACTTTAAACCAACACCGCTGAAGGTGAACCTCGGAACTTTGTTTCATAAAATCTTTCAAA	2786
DB	3992	TCCACTTTAAACCAACACCGCTGAAGGTGAACCTCGGAACTTTGTTTCATAAAATCTTTCAAA	4051
QY	2787	AGTTGTTTTTATCATCAATGTTTAAATTTCAAAATGCTGCAGGGTAAATTTAATGTATATAAAT	2846
DB	4052	AGTTGTTTTTATCATCAATGTTTAAATTTCAAAATGCTGCAGGGTAAATTTAATGTATATAAAT	4111
QY	2847	ATTAGTAAAGAAAGATGATGATTCACATCTTAGTATGAAATAGATACAAACATACAAATTC	2906
DB	4112	ATTAGTAAAGAAAGATGATGATTCACATCTTAGTATGAAATAGATACAAACATACAAATTC	4171
QY	2907	ATTCAAGTGCATGCTTTAGGTGTTTAAAGCATGAGATTTGTACATGTTTACTGTTAGGTCCTTG	2966
DB	4172	ATTCAAGTGCATGCTTTAGGTGTTTAAAGCATGAGATTTGTACATGTTTACTGTTAGGTCCTTG	4231
QY	2967	CATCTGTGCTGCTAGGTGAGTATGAGAAGATGTCAGGAATGTCGACGTATTTGTTGCTTA	3026
DB	4232	CATCTGTGCTGCTAGGTGAGTATGAGAAGATGTCGAGGAATGTCGACGTATTTGTTGCTTA	4291
QY	3027	AAAAAAAAGGCTGTTTGTAGGCGTTTAAATATGCTTATTTGTGCTGCTCACACTACC	3086
DB	4292	AAAAAAAAGGCTGTTTGTAGGCGTTTAAATATGCTTATTTGTGCTGCTCACACTACC	4351
QY	3087	TATTACACACTGTTGCTTGTGCGGTTTGTGTTGTTGATGCGGTGTTATACAGTAGTTAA	3146
DB	4352	TATTACACACTGTTGCTTGTGCGGTTTGTGTTGTTGATGCGGTGTTATACAGTAGTTAA	4411
QY	3147	ATTTCCTCATGAGAAAAATAAATGTCCTGAATTTCTCAA	3184
DB	4412	ATTTCCTCATGAGAAAAATAAATGTCCTGAATTTCTCAT	4449
RESULT 9			
AX127525		2738 bp	linear
LOCUS	AX127525	2738 bp	DNA
DEFINITION	Sequence 1 from Patent WO0131007.		PAT 15-MAY-2001
ACCESSION	AX127525		
VERSION	AX127525.1	GI:14134230	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS	Chiang, L.W.		
TITLE	Nucleic acid molecules derived from rat brain and programmed cell death models		
JOURNAL	Patent: WO 0131007-A 1 03-MAY-2001;		
	Millennium Pharmaceutical, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1. .2738		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		

RESULT	9
AXI27525	
LOCUS	DNA linear PAT 15-MAY-2001
DEFINITION	Sequence 1 from Patent WO0131007.
ACCESSION	AXI27525
VERSION	AXI27525.1 GI:14134230
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1
REFERENCE	Chiang, L.W. Nucleic acid molecules derived from rat brain and programmed cell death models
TITLE	Patent: WO 0131007-A 1 03-MAY-2001;
JOURNAL	Millennium Pharmaceuticals, Inc. (US) Location/Qualifiers
FEATURES	1..2738 <code>/organism=</code> "Homo sapiens" <code>/mol_type=</code> "unassigned DNA" <code>/db_xref=</code> "taxon:9606"
SOURCE	

Query Match	77.4%	Score 2481.8	DB 6	Length 2738
Best Local Similarity	99.7%	Pred. No. 0		
Matches 2486	Conservative	0	Mismatches	7
			Indels	0
			Gaps	0
Qy	692	CCAAATAGCTGGAGATATCTTTAATAGCGCAATAGATTCAGTGCAGGCATTTCACAGC	751	
Db	7	CCAGCGCTCGGAGATATCTTTAATAGCGCAATAGATTCAGTGCAGGCATTTCACAGC	66	

QY 752 CGGAGTGTGGTTATGGCTTGAGCCTGATCGTTGGACAGAGTACAGATACAGACGATGG 811
Db |||||
QY 67 CGGAGTGTGGTTATGGCTTGAGCCTGATCGTTGGACAGAGTACAGATACAGACGATGG 126
Db |||||
QY 812 AACACAGATAACCTGGAACTAATCTTTTCATTTTTTCGAAGAAGATCTCAGTGGACAGCTAG 871
Db |||||
QY 127 AACACAGATAACCTGGAACTAATCTTTGATTTTTTCGAAGAAGATCTCAGTGGACAGCTAG 186
Db |||||
QY 872 TTCAGGGTGATGCCCTTCTCGACATGTGGGTACAGCTTGTCTCTTATCATCATCCACCATTTG 931
Db |||||
QY 187 TTCAGGGTGATGCCCTTCTCGACATGTGGGTACAGCTTGTCTCTTATCATCATCCACCATTTG 246
Db |||||
QY 932 CTGAGAGTGGAAAGAGTGTGGAAATCTTACTCTTCCCATCATGAGCAGAGAAATTCGCGGA 991
Db |||||
QY 247 CTGAGAGTGGAAAGAGTGTGGAAATCTTACTCTTCCCATCATGAGCAGAGAAATTCGCGGA 306
Db |||||
QY 992 AAAAATAGGCAAGTGTGAGAGTGTGACTATATAATTTAAGCCATTAACAGGATACAGTT 1051
Db |||||
QY 307 AAAAATAGGCAAGTGTGAGAGTGTGACTATATAATTTAAGCCATTAACAGGATACAGTT 366
Db |||||
QY 1052 GTGACATGAATCTTCAATTTTCCAAAGTATTGGAAGCCAAAGAAATACCAATTTGGATTTGGCC 1111
Db |||||
QY 367 GTGACATGAATCTTCAATTTTCCAAAGTATTGGAAGCCAAAGAAATACCAATTTGGATTTGGCC 426
Db |||||
QY 1112 ATCGAGGTGCAGGAACTCTACAACTGCCAGCTGGCTTAAAGTTTCAGAGAAATACTA 1171
Db |||||
QY 427 ATCGAGGTGCAGGAACTCTACAACTGCCAGCTGGCTTAAAGTTTCAGAGAAATACTA 486
Db |||||
QY 1172 TTGCTTTTAAAGAAATGCTGTAGTGTGAGAGCCTTTGTAGAAATTTGACGTACACC 1231
Db |||||
QY 487 TTGCTTTTAAAGAAATGCTGTAGTGTGAGAGCCTTTGTAGAAATTTGACGTACACC 546
Db |||||
QY 1232 TTTTAAAGAGCTTTGTGCCGTGGTATATCATGATCTTTACCTGTGTTTGTGACTATGAAAA 1291
Db |||||
QY 547 TTTTAAAGAGCTTTGTGCCGTGGTATATCATGATCTTTACCTGTGTTTGTGACTATGAAAA 606
Db |||||
QY 1292 AGAAATTTGATGCTGATCCAGTTGAAATTTATTTGAAATTTCCAGTAAAGAAATTAACATTTG 1351
Db |||||
QY 607 AGAAATTTGATGCTGATCCAGTTGAAATTTATTTGAAATTTCCAGTAAAGAAATTAACATTTG 666
Db |||||
QY 1352 ACCAATCCAGTTGTTTAAAGCTCACTCATGTGACCTGCACTGAAATCTAAGGATCGGAAAG 1411
Db |||||
QY 667 ACCAATCCAGTTGTTTAAAGCTCACTCATGTGACCTGCACTGAAATCTAAGGATCGGAAAG 726
Db |||||
QY 1412 AATCTGTGTTTCAGAGGAGAAATTCCTTTTCAGAAAAATCAGCAATTCCTTCTCTTAAGA 1471
Db |||||
QY 727 AATCTGTGTTTCAGAGGAGAAATTCCTTTTCAGAAAAATCAGCAATTCCTTCTCTTAAGA 786
Db |||||
QY 1472 TGGTTTTAGAGTCTTTGCCAGAGATGTAGGGTTTAACTTGAATTAATAATGGATCTGCC 1531
Db |||||
QY 787 TGGTTTTAGAGTCTTTGCCAGAGATGTAGGGTTTAACTTGAATTAATAATGGATCTGCC 846
Db |||||
QY 1532 AGCAAGGATGGAATGTGGATGGTAACTTATCAACATATTTTGACATGATCTGTTTT 1591
Db |||||
QY 847 AGCAAGGATGGAATGTGGATGGTAACTTATCAACATATTTTGACATGATCTGTTTT 906
Db |||||
QY 1592 TGGATATAATTTTAAAAATCTTTTAAAGAAATTTCTGGGAAGAGGAAATAGTGTGTTTTCTT 1651
Db |||||
QY 907 TGGATATAATTTTAAAAATCTGTTTTTAAAGAAATTTCTGGGAAGAGGAAATAGTGTGTTTTCTT 966
Db |||||
QY 1652 CATTTGATGAGATATTTTGACAAATGGTTGCGCAAAAGAGAGAAATAATTCGATACTAT 1711
Db |||||
QY 967 CATTTGATGAGATATTTTGACAAATGGTTGCGCAAAAGAGAGAAATAATTCGATACTAT 1026
Db |||||
QY 1712 TTTTAACTCAAGGAAATCTGAGATTTATCTTGAATCTGAGACCTCAGATCTCGGACAA 1771
Db |||||
QY 1027 TTTTAACTCAAGGAAATCTGAGATTTATCTTGAATCTGAGACCTCAGATCTCGGACAA 1086
Db |||||
QY 1772 CCCCCATTTGCAATGTAGCTTTGCAAGTTTCAAAATCTTACTGGGGATAAATGTACATACTG 1831
Db |||||
QY 1087 CCCCCATTTGCAATGTAGCTTTGCAAGTTTCAAAATCTTACTGGGGATAAATGTACATACTG 1146
Db |||||

QY 1832 AAGACTTCTCAGAAAAACCCATCTATATTCAAGAGGCAAAAGCTTAAGGACTAGTCATAT 1891
Db |||||
QY 1147 AAGACTTCTCAGAAAAACCCATCTATATTCAAGAGGCAAAAGCTTAAGGACTAGTCATAT 1206
Db |||||
QY 1892 TCTGCTGGGGTGATGATACCAATGATCTGGAAGAAAGAGAAATTTGAAGAACTTGGAG 1951
Db |||||
QY 1207 TCTGCTGGGGTGATGATACCAATGATCTGGAAGAAAGAGAAATTTGAAGAACTTGGAG 1266
Db |||||
QY 1952 TTAATGGTCTAATTTATGATAGGATATATGATGGATGCTGAAACCAACCAATATATTTCC 2011
Db |||||
QY 1267 TTAATGGTCTAATTTATGATAGGATATATGATGGATGCTGAAACCAACCAATATATTTCC 1326
Db |||||
QY 2012 AAGTGGAGCAATTTGAACCGCTGAAGAGGAATTTGCCAGAGCTTTAAGAGCTGTTTGTGTC 2071
Db |||||
QY 1327 AAGTGGAGCAATTTGAACCGCTGAAGAGGAATTTGCCAGAGCTTTAAGAGCTGTTTGTGTC 1386
Db |||||
QY 2072 CCACTGTTAGCCGCTTTGTTCCCTCATCTTTGTTGGGAGTCTGATATCATATGTTGGATG 2131
Db |||||
QY 1387 CCACTGTTAGCCGCTTTGTTCCCTCATCTTTGTTGGGAGTCTGATATCATATGTTGGATG 1446
Db |||||
QY 2132 CCAACGGCAATTTGAACCGCTGAAGAGGAATTTGCCAGAGCTTTAAGAGCTGTTTGTGTC 2191
Db |||||
QY 1447 CCAACGGCAATTTGAACCGCTGAAGAGGAATTTGCCAGAGCTTTAAGAGCTGTTTGTGTC 1506
Db |||||
QY 2192 GCGTGCACCGCTGTTCTGGGTATTCAATTTTCACTGAGCAATTTGTTGATCTATGCCTT 2251
Db |||||
QY 1507 GCGTGCACCGCTGTTCTGGGTATTCAATTTTCACTGAGCAATTTGTTGATCTATGCCTT 1566
Db |||||
QY 2252 TTGGGGTCTCAGTTCAATGAAGCAATTAATGAAGTATTTAACTCTTTCACTACAGTTCTT 2311
Db |||||
QY 1567 TTGGGGTCTCAGTTCAATGAAGCAATTAATGAAGTATTTAACTCTTTCACTACAGTTCTT 1626
Db |||||
QY 2312 GCAAGTATGCTATTTAAATTTACTTTGCCAGGTATTAATTTGCCAGTCACTCTCTTTATAGTG 2371
Db |||||
QY 1627 GCAAGTATGCTATTTAAATTTACTTTGCCAGGTATTAATTTGCCAGTCACTCTCTTTATAGTG 1686
Db |||||
QY 2372 AGAAAAATTTATTTGGTTAGTAAATATAAATTTTAACTAAATATAATAATCTATAATGTT 2431
Db |||||
QY 1687 AGAAAAATTTATTTGGTTAGTAAATATAAATTTTAACTAAATATAATAATCTATAATGTT 1746
Db |||||
QY 2432 AAAAATATGTTTCAATTAAGCAATGACCTTTGAAATTAACATATAATAATAGCTCATATTT 2491
Db |||||
QY 1747 AAAAATATGTTTCAATTAAGCAATGACCTTTGAAATTAACATATAATAATAGCTCATATTT 1806
Db |||||
QY 2492 ACATTTACAGCTTTTCAATTTGATCAGGTCTGAAATCTTTAGCACTTTAAGGAAATGACATA 2551
Db |||||
QY 1807 ACATTTACAGCTTTTCAATTTGATCAGGTCTGAAATCTTTAGCACTTTAAGGAAATGACATA 1866
Db |||||
QY 2552 TGCATTAATTAACCTGACCAATGAAAAATAAGTACCTGAAATGCAATTTGCACTGG 2611
Db |||||
QY 1867 TGCATTAATTAACCTGACCAATGAAAAATAAGTACCTGAAATGCAATTTGCACTGG 1926
Db |||||
QY 2612 TGATTTCAACTGCACAAAATCTTTTGCCATCTTTGATATAGGTATTTTACATGGTTG 2671
Db |||||
QY 1927 TGATTTCAACTGCACAAAATCTTTTGCCATCTTTGATATAGGTATTTTACATGGTTG 1986
Db |||||
QY 2672 ACATGCAACAACCACTTTTCAATTTAGTATGAACCTTTGAGGCTGCTGCCATTTTCCAC 2731
Db |||||
QY 1987 ACATGCAACAACCACTTTTCAATTTAGTATGAACCTTTGAGGCTGCTGCCATTTTCCAC 2046
Db |||||
QY 2732 TTAACCAAAACCCAGCTGAAGGTGAACCTCGAAACCTGTTTTCATATAATCTTTCAAAAGTTG 2791
Db |||||
QY 2047 TTAACCAAAACCCAGCTGAAGGTGAACCTCGAAACCTGTTTTCATATAATCTTTCAAAAGTTG 2106
Db |||||
QY 2792 TTTTACATCAATGTTTAAAAATTTTCAAAATGCTGCAAGGTAAATTTAATCTATAATAATATTAG 2851
Db |||||
QY 2107 TTTTACATCAATGTTTAAAAATTTTCAAAATGCTGCAAGGTAAATTTAATGATATAATAATATTAG 2166
Db |||||
QY 2852 TAAGAAAAAGTATGTTTGGATCTTTAGTAGAATAGATCAACAATACAAATTTCAATTTCA 2911
Db |||||
QY 2167 TAAGAAAAAGTATGTTTGGATCTTTAGTAGAATAGATCAACAATACAAATTTCAATTTCA 2226
Db |||||
QY 2912 GTGCATGCTTTTAGTGTGTTAAGCATGAGATGTAATGTTTACTGTTAGGTCTTGCATCT 2971
Db |||||

```
|||||
2227 GTGCATGCTTTAGTGTAAAGCATGAGATGTACATGTTACTGTAGTCTTGCATCT 2286
QY 2972 GTGGTCTAGGTGAGTATGAGAAGATGTCAAGGACTGGACGATTTTGTGCTTAAAAA 3031
Db 2287 GTGGTCTAGGTGAGTATGAGAAGATGTCAAGGACTGGACGATTTTGTGCTTAAAAA 2346
QY 3032 AAAAGCTGTTGTAGCGCTTTTAAATATGCTTATTTTGTGCTCTCACTACCTATTA 3091
Db 2347 AAAAGCTGTTGTAGCGCTTTTAAATATGCTTATTTTGTGCTCTCACTACCTATTA 2406
QY 3092 CACACTGTTGCTTTGGGCTTTGTTTGTATGCGGTGTTATACAGTAGTAAATTTTC 3151
Db 2407 CACACTGTTGCTTTGGGCTTTGTTTGTATGCGGTGTTATACAGTAGTAAATTTTC 2466
QY 3152 CATGCAGAAAAATATATGCTCCTGAATTTCTCAAA 3184
Db 2467 CATGCAGAAAAATATATGCTCCTGAATTTCTATA 2499

RESULT 10
AC144439
LOCUS 39694 bp DNA linear HTG 16-APR-2003
DEFINITION Homo sapiens chromosome 7 clone XFOS-87042G10, WORKING DRAFT
AC144439
VERSION 1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston.R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 39694)
Waterston.R.H.
Direct Submission
Submitted (16-APR-2003) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: H_AA087042G10
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 39167 bases at least Q40
Consensus quality: 39319 bases at least Q30
Consensus quality: 39366 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 7028: contig of 7028 bp in length
* 7029 7128: gap of unknown length
* 7129 16822: contig of 9694 bp in length
* 16823 16922: gap of unknown length
```

```
FEATURES
source
* 16923 39694: contig of 22772 bp in length.
Location/Qualifiers
1..39694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/clone="XFOS-87042G10"
1..7028
/misc_feature
/note="assembly_name:Contig4"
7129..16822
/misc_feature
/note="assembly_name:Contig5"
16923..39694
/misc_feature
/note="assembly_name:Contig6"

ORIGIN
Query Match 72.7%; Score 2329.4; DB 2; Length 39694;
Best Local Similarity 84.2%; Pred. No. 0;
Matches 2993; Conservative 0; Mismatches 141; Indels 419; Gaps 16;

QY 11 GCCTCCGGCGGAGGACGACGCGGCGCGCGGTACCTCTGCGCGCGGTCTCTCTCTCTC 70
Db 32322 GAGCGCTGCGGAGGACCGACGAGCGGGGTCCGATACCTCTGCGCGGTACTCTCTCTC 32381
QY 71 GGGCGGGCGGCGGACGCGGACCTGCGGACTAGCGAACCCGG--AGCACGACATCATAA 129
Db 32382 GGGCGG---GGCGGCAACGCGGACCCGTGGACTAGCGAACCCGGCAGCAGCATCATAA 32438
QY 130 AATAATCCATCAGAAATGACACCTTCTCAGGTGCGCTTTGAAATAGAGGACTCTTTA 189
Db 32439 AATAAACCCATCAGAAATGACACCTTCTCAGGTGCGCTTTGAAATAGAGGACTCTTTA 32498
QY 190 CCAGGAGAGGTTTTTGGCGATATGTGGAAGCTGTGATGCTTTGGGAACTTGAATCCTCAA 249
Db 32499 CC---AGAAGTTTTTGCAATATGTGGAAGCTGTGATGCTTTGGGAACTTGAATCCTCAA 32555
QY 250 AATGCTGTGGCTCTTTCCAGAGAAATGACACAGGTGAA----- 288
Db 32556 AATGCTGTGGCTCTTTCCAGAGAAATGACACAGGTGGAAGAACTGATGATCAGA 32615
QY 289 -----AGCATGCTATCGAAGCAACCAT 311
Db 32616 ATGAGGGCAATGCCCTAAAACTGTCTTAGCAGAGCCAGCATGTTATGGAAGCAACCAT 32675
QY 312 TGTACTCAGTAGAGGAGTATCAGTTTCAGTATCGCTACTTCAAAGGGTACTTTTGAAGACC 371
Db 32676 TGTACTCAGTAGAGGAGTATCAGTTTCAGTATCGCTACTTCAAAGGGTACTTTTGAAGACC 32735
QY 372 AAAGACTATCGGTGGTCCATGTCAAGTATGTTTCAAGTGGGAGACTCATCTTCAACC 431
Db 32736 AAAGACTATCGGTGGTCCATGTCAAGTATGTTTCAAGTGGGAGACTCATCTTCAACC 32795
QY 432 ACGATCAATAACCCCTTTAGAAAGGGAATTTATTATGACGATGACAAATTTGGGAATCCA 491
Db 32796 ACGATCAATAACCCCTTTAGAAAGTGAATTTGTTATTGACCATGGAACAATTTGGGAATCCA 32855
QY 492 CAATGGTGTGAAACTCTGGATTCTGGATCGCTGACATGTCAAGTGAATAAGATTACG 551
Db 32856 CAGTGGTGTGAAACTCTGGATTCTGGATCGCTGACATGTCAAGTGAATAAGATTACG 32915
QY 552 TTTGCATTATTTCTGAAAAACCTCTGCTCAATACCAAGAAAAAATTTAAAAAATCTAG 611
Db 32916 ATTGCATTATTTCTGAAAAACCTCTGCTCAATACCAAGAAAAAATTTAAAAAATCTAG 32975
QY 612 ATTTAGGTGAGCTGACACTAGAGGCTGGAGGAATGACGATGATAGGTATCTCC 671
Db 32976 ATTTAGGTGAGCTGACACTAGAGGCTGGAGGAATGATAGGTATCTCC 33035
QY 672 CACTGTACTCCAAAAATGTCCAATAGCTTGGAGATATCTTTAATAAGCGACAATGAGTT 731
Db 33036 CACTGTACTCCAAAAATGTCCAATAGCTTGGAGATATCTTTAATAAGCGACAATGAGTT 33095
QY 732 CAAGTGCAGGCAATTCACAGCCGGAGTGTGGTTATGGCTTGCAGCCTGATCGTTGGACAGA 791
```

Db	33096		CCAGTCGAGGCATTCACAGCTGGAGTGGTATGGCTTACAGTCTGATCGTTGGACAGA	33155	Db	34165	TAGGATAGGCTGGGTGGTGGCTCGCGCTGTAAACCCACAGACTTTGGGAGCCGAGG	34224
Qy	792		GTACAGCATACAGACGATGGAACACAGATAACCTGGAACTAATCTTTGATTTTTTCGAAGA	851	Qy	1816	-----	1815
Db	33156		GTACAGCATACAGACAATGGACACAGATACCTGGAACTAATCTTTGATTTTTTCGAAGA	33215	Db	34225	AGGTGGATCACCTGAGGTCAGAAAGTTCAAGATCAGCTTAGTCAACATGGTGAACCCAG	34284
Qy	852		AGATCTCAGTCAGCAGCAGTGTTCAGGGTGATGCCCTTCCTGGACATGTGGGTACAGCTTG	911	Qy	1816	-----	1815
Db	33216		AGATCTCAGTCAGCAGCAGTGTTCAGGGTGATGCCCTTCCTGGATGTGGGTACAGCTTG	33275	Db	34285	TCTCTACTAATATACAAAAATTAGCTGGGCACCGTGGTGGCACCCTGTATATCCAGCTA	34344
Qy	912		TCTCTTATCATCCACCATGCTGAGAGTGGAAAGAGTGGTGGATCTTACTCTCCCAT	971	Qy	1816	-----	1815
Db	33276		TCTCTTATCATCCACCATGCTGAGAGTGGAAAGAGTGGTGGATCTTACTCTCCCAT	33335	Db	34345	CTGGCAGGCTGAGCAGGAGAAATCGCTTGAATCCGGGAGGCGGAGGTTGCAGTGAGTCG	34404
Qy	972		CATGAGCAGAAATTCGCGGAAAAACAATAGGCAAGTGGAGTGGATCATATATATTTAA	1031	Qy	1816	-----	1815
Db	33336		CATGAGCAGAAATTCGAGGAAAAACAATAGGCAAGTGGAGTGGATCATATATATTTAA	33395	Db	34405	AGATCACACCAATTTGTCTCCGGCTGGGAAACAAGAGGAAACCTTCGTCTAAAAAATAAA	34464
Qy	1032		GCCATTACAGGATACAGTGTGACATGAATCTTCAATTTTCCAAATTTGGAGCCCAAG	1091	Qy	1816	-----ATAAATGTACATACTGAAGACTTCTCAGAAACCCA	1851
Db	33396		GCCATTGCGAGATACAGTGTGACATGAATCTTCAATTTTCCAAATTTGGAGCCCAAG	33454	Db	34465	GAAAGAAAGAAAAATCTACTGGAATTAATGTACATACTGAAGACTTCTCAGAAACCCA	34524
Qy	1092		AATACCATTTGGATGTTGCCATCGAGGTGCAAGGAAAATCTACACAACTGCCAGCTGGC	1151	Qy	1852	TCTATATTTCAAGAGGCAAAAAGCTTAAGGACTAGTGCATATTTCTGCTGGGTGATGATACC	1911
Db	33455		AATACCATTTGGATGTTGCCATCGAGGTGCGGAAACGCT---ACAACCTGCCAGCTGGC	33511	Db	34525	TCTATATTTCAAGAGGCAAAAAGCTTAAGGACTAGTGCATATTTCTGCTGGGTGATGAAAGC	34584
Qy	1152		TAAAGTTCAAGAAATACATTTGCTTCTTAAAGAAATGCTCTAGTCAATGGTGACGCTT	1211	Qy	1912	AATGATCTCTGAAAACAGAGGAAATTTGAGGAACCTTGGAGTTAATGGTCTAATTTATGAT	1971
Db	33512		TA-----AAGAAATACATTTGCTTCTTAAAGAAATGCTCTAGTCAATGGTGACGCTT	33565	Db	34585	AGTGATCTCTGAAAAGGAAAGAAATTTGATGGAACCTTGGAGTTAATGGTCTAATTTATGAT	34644
Qy	1212		TCGTAGAAATTTGACGTACACCTTTCAAAGGACTTTGTGCCGTGGTATATCATGATCTTAC	1271	Qy	1972	AGGATATATGATTTGGATGCTGAAACAAACCAATATATTTCCAAAGTGGAGCAATTTGAAAGC	2031
Db	33566		TGTAGAAATTTGATGTACACCTTTCAAAGGACTTTGTGCCATGTATATCATGATCTTAC	33625	Db	34645	AGGATATATGATTTGGATGCTGAAACCA-----CCAAGTGGAGCACTGGAATGC	34693
Qy	1272		CTGTTGTTTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTTGAATTC	1331	Qy	2032	CTGAACAGGAAATGCGCAGAGCTTAAAGACTGTTTGTGTCCCACTGTTAGCCGCTTTGTT	2091
Db	33626		CTGTTGTTTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTTGAATTC	33685	Db	34694	CTGAACAGGAAATGCGCAGAGTGAAGGGCTGTTTGTGTCCCACTGTTAGCCACCTTTGTT	34753
Qy	1332		AGT-AAAAGAAATTAACATTTGACCACTCCAGTGTGTTAAAGCTCACTGATGATCGAC	1390	Qy	2092	CCCTCATCTTTGTGGGGAGTCTGATATCCATGTGGATGCCAATGGCAATGATATGATG	2151
Db	33686		AGTAAAAAAATTAACATTTGACCACTCCAGTGTGTTAAAGCTCACTGATGATCGAC	33745	Db	34754	CCCTCATCTTTGTGGAGGCTGATACCCATGTGGATGCCAATGGCAATGATATGATG	34813
Qy	1391		TCGAATCTAAGATCGGAAAGAAATCTGTGTTTCAGGAGGAAAAATTCCTTTTCAGAAAAATC	1450	Qy	2152	GAGAACTCTAGTTTTTATTTGCACAGAGTCAATTTTGGGGGGTGCACCGCTGTTCTGGG	2211
Db	33746		TGAATCTAAGATTTGAAAGAAATCTGTGTCCAGGAGGAAAAATTCCTTTTCAGAAAAATC	33805	Db	34814	GAGAACTCTAGTTTTTATTTGCACAGAGTCAATTTTGGGGGCAATGACCGCTGTTCTGGG	34873
Qy	1451		AGCCATTTCTCTTTAAGATGGTTTTAGAGTCTTTTGCAGAGATGTAGGGTTTAAACA	1510	Qy	2212	TATTCATTTTTCATCACTGAGCAATGTTGATCTATGCTTTTGGGGCTTCTCAGTTCAATG	2271
Db	33806		AGCCATTTCTCTCTGTAAGATGGTTTTAGAGTGTTTTGCAGAGATGTAGGGTTTAAACA	33865	Db	34874	TATTCATTTTTCATCACTGAGCAATGTTGATCTATGCTTTTGGGGCTTCTCATTCAACG	34933
Qy	1511		TTGAAATAAATGGATCTGCCAGCAAGGATGGATGGGATGTAACCTTATCAACAT	1570	Qy	2272	AAGCAATAATGAAGTATTTAACTCTTTCACACAGTCTTTCGCAAGTATGCTATTTAAAT	2331
Db	33866		TTGAAATAAATGGATCTGCCAGCAAGGATGGATGGGATGTAACCTTATCAACAT	33925	Db	34934	AAACAATAATGAAGTATTTAACTCTTTCACATAGTCTTTCGCAAGTATGCTATTTAAAT	34993
Qy	1571		ATTTTCACATGATCTGTTTTTGGATATAATTTTAAAAATCTGTTTAGAAAAATTCGGGA	1630	Qy	2332	ACTTTGGCCAGTATAATTTGCCAGTCACTCTTTTATAGTAGAGAAAAATTTATGTTAGTA	2391
Db	33926		ATTTTGGCATGAATCTGTTTTTGGATATAACTTTTAAAAATCTGTTTAGAAAAATTCGGGA	33985	Db	34994	ACTTAGCCAGGTGAATTTGCCAGTCACTCTTTACAGTGAAGAAATGATTTGTTGTA	35053
Qy	1631		AGAGGAGAAATGATGTTTTCTTCATTTGATGACAGATATTTGCAATGTTTCGGCAAGGC	1690	Qy	2392	ATATAAATATTTTAAACTTAAATATATAAATCTATAATGTTTAAACATATGTTCAATTAAG	2451
Db	33986		AGA-GAGAAATGATGTTTTCTTCATTTGATGACAGATATTTTGCACAAATGTTTCGGCAAGGC	34044	Db	35054	ATATAAATATTTTAAACTTAAATATATAAATCTATAATGTTTAA-----ATGTTCTTAAAG	35109
Qy	1691		AGAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGAACTCA	1750	Qy	2452	CATAGCACTTTGAAATTAACATATAATAAATAGCTCATATTTACACTTACAGCTTTTCAAT	2511
Db	34045		AGAACAAATATCCCATATTTATTTTAACTGAAGGAAATCTGAGATTTATCTGAACTCA	34104	Db	35110	CATAGCACTTTGAAATTAATAATAAATAGCTCATATAAATAGCTCATATAAATAGCTTTT	35168
Qy	1751		TGACCTCAGATCTCGACAAACCCCATTTGCAATGAGCTTTGACAGTTTGAATTTCTAC	1810	Qy	2512	GATCAGGCTGAAATCTTTAGCACTTAAGGAAATGACTATGCTAATAATATACCTGACCA	2571
Db	34105		TGACCTCAGATCTCGACAAACCCCATTTGCAATGAGCTTTGACAGTTTGAATTTCTAC	34164	Db	35169	GATCAGGCTGAAATCTTTAGCACTTTAAGGAAATGACTATGCTAATAATATACCTGACCA	35228
Qy	1811		TTGGGG-----	1815	Qy	2572	TGAAAAAATAAGTACCTTCAAAATGCAATTTGCACTGGTGAATTTCAACTGCAAAATC	2631
					Db	35229	TGAAAAAATAAGTACCTTCAAAATGCAATTTGCACTGGTGAATTTCAACTGCAAAATC	35288

11 GCGTCCGGCGGAGGACGGACGGCGCGGTACCTCTGCCCCGGGTCTCGGTCTC

Db 133682 GAGCGTGGCGAGGACCGACAGCGGGGTCCGGTACCTCTCCCGCGGTCTCTCTCTC 133623
Qy 71 GGGCGGGCGGGGCGACCGCGACCTGCGGACTAGCGAACCCGG-AGCACGACATCAAA 129
Db 133622 GGGCGG---GGCGGCGACCGGACCGGTGAACCTAGCGACCCCGGACGACATCAAA 133566
Qy 130 AATAAATCATCAGAAATGACACTCTCTCAGGTGCTTTGAAATTAAGAGGAACTCTTTTA 189
Db 133565 AATAAACCCATCAGAATGACACTCTCTCAGGTGCTTTGAAATTAAGAGGAACTCTTTTA 133506
Qy 190 CCAGGAGAGTGTTCGATATGCGAAGCTGTGATGCTTTGCGAACTTGGAAATCCTCAA 249
Db 133505 CC---AGAAAGTTTTGCAATATGTGAAAGCTGTGATGCTTTGCGAACTTGGAAATCCTCAA 133449
Qy 250 AATGCTGTGGCTCTTCTCCAGAGAAATGACACAGGTGAA-----288
Db 133448 AATGCTGTGGCTCTTCTCCAGAGAAATGACACAGGTGGAAGCTAAGATTGTATCAGA 133389
Qy 289 -----AGCATGCTATGGAAGCAACCAT 311
Db 133388 ATGAGGGCAATGCCCTTAAACCTGTCTTAGCAGAGCCAGCATGTTATGGAAGCAACCAT 133329
Qy 312 TGTACTCAGTAGAGGATATCAGTTTCAGTATCGTACTTCAAGGGTACTTTTGAAC 371
Db 133328 TTTACTCAGTAGAGGATATCAGTTTCAGTATCGTACTTCAAGGGTACTTTTGAAC 133269
Qy 372 AAAGACTATCGGTGGTCCATGTCAAGTGATAGTTCAAGTGGGAGACTCATCTACAAC 431
Db 133268 AAAGACTATCGGTGGTCCATGTCAAGTGATAGTTCAAGTGGGAGATTCATCTACAAC 133209
Qy 432 AGCATCAATPAAACCCCTTTAGAAAGGCAATTAATTTGACGATGGACAAATTTGGAATCA 491
Db 133208 AGCATCAATPAAACCCCTCAGAAAGTGAATTTGTTATGACCATGGACAAATTTGGAATCA 133149
Qy 492 CAATGGTGTGAACTCTGGAATCTGGATGGCTGACATGCTGACATGAAATTAAGATTACG 551
Db 133148 CAGTGGTGTGAACTCTGGAATCTGGATGGCTGACATGCTGACATGAAATTAAGATTACG 133089
Qy 552 TTTGCAATATTCTGAAACCTCTGTGTCAATTAACCAAGAAATAATTAATAAATCTAG 611
Db 133088 TTTGCAATATTCTGAAACCTCTGTGTCAATTAACCAAGAAATAATTAATAAATCTAG 133029
Qy 612 ATTTAGGTGAACTGACACTAGAGGCTGAGGAAAGATGACGATAGGGTATCTCC 671
Db 133028 ATTTAGGTGAACTGACACTAGAGGCTGAGGAAAGATGACGATAGGGTATCTCC 132969
Qy 672 CACTGTACTCCAAAATGTCGAATAGCTTGAGATATCCTTAATAAGGCAATAGATT 731
Db 132968 CACTGTACTCCAAAATGTCGAATAGCTTGAGATATCCTTAATAAGGCAATAGATT 132909
Qy 732 CAAGTGCAGGCATTCACAGCCGAGTGTGTTATGGCTTGACGCTGATCGTTGGACAGA 791
Db 132908 CCAGTGCAGGCATTCACAGCTGAGTGTGTTATGGCTTGACGCTGATCGTTGGACAGA 132849
Qy 792 GTACAGCATACAGACGATGAACCGATTAACCTGGAACTAATCTTTGATTTTTCGAAGA 851
Db 132848 GTACAGCATACAGACGATGAACCGATTAACCTGGAACTAATCTTTGATTTTTCGAAGA 132789
Qy 852 AGATCTCAGTAGCAGGTATGTTACGGGTGATGCCCTTCTCGACATGTTGGGTACAGCTTG 911
Db 132788 AGATCTCAGTAGCAGCATATGTTCAAGGTGATGCCCTTCTCGACGTTGGGTACAGCTTG 132729
Qy 912 TCTCTTATCATCCACCATTTGCTGAGATGGAAGAGTGTGGAATTTCTTACTCTTCCCAT 971
Db 132728 TCTCTTATTTTCAACCATTTGCTGAGATGGAAGAGTGTGGAATTTCTTACTCTTCCCAT 132669
Qy 972 CATGAGCAGAAATTCGCGAAACCAATAGCAGAGTGTGATATTAATTTAA 1031
Db 132668 CATGAGCAGAAATTCGCGAAACCAATAGCAGAGTGTGATATTAATTTAA 132609
Qy 1032 GCCATTACAGGATACAGTTGTGACATGAATCTTCAATTTTCCAAAGTATGGAAGCCAAAG 1091
Db 132608 GCCATTGCGAGATACAGTTGTGACATGAATCTTCAATTTTCC-AGTATTGGAAGCCAAAG 132550

Qy 1092 AATACCATTTGGATGTTGCCATCGAGGTGACGAAACTCTTACAACTGCCAGCTGGC 1151
Db 132549 AATACCATTTGGATGTTGCCATCGAGGTGTCGGAAC---ACTACAACTGCCAGCTGGC 132493
Qy 1152 TAAAGTTCAAGAAATACTATTGCTCTTTTAAAGAAATGCTCTAGTCATGTCGACGCTT 1211
Db 132492 TA-----AAGAAATACTATTGCTCTTTTAAAGAAATGCTCTAGTCACAGTGCAGCTT 132439
Qy 1212 TGTAGAATTTGACGTACACCTTTCAAAGGACTTTGTGCCCCGTGGGTATATCATGATCTTAC 1271
Db 132438 TGTAGAATTTGATGTACACCTTTCAAAGGACTTTGTGCCCCGTGGGTATATCATGATCTTAC 132379
Qy 1272 CTGTTGTTGATATGAAAAAGAAATTTGATGCTGATCCAGTCCAGTGAATTTTGAATTTCC 1331
Db 132378 CTGTTGTTGATATGAAAAAGAAATTTGATGCTGATCCACTTGAAATTTTGAATTTCC 132319
Qy 1332 AGT-AAAAGAAATTAACATTTGACCAACTCTGAGTTGTTTAAAGCTCCTCATGCTGAC 1390
Db 132318 AGTAAAAAAATTAACATTTGACCAACTCTGAGTTGTTTAAAGCTCCTCATGCTGAC 132259
Qy 1391 TGAATCTAAGGATCGGAAAGAAATCTGTGTTTCAGGAGGAAATTTCTTTTCAGAAAAATC 1450
Db 132258 TGAATCTAAGGATTTGAAACAACTCTGTGTTCCAGGAGGAAATTTCTTTTCAGAAAAATC 132199
Qy 1451 AGCCATTTCTCTCTTAAAGATGGTTTTAGAGTCTTTGCCAGAAAGATGTAGGGTTTAA 1510
Db 132198 AGCCATTTCTCTCTCGTAAGATGGTTTTAGAGTGTGTTGCCAGAAAGATGTAGGGTTTAA 132139
Qy 1511 TTGAATTAATTAATGATCTGCCAGCAAGGATGGAATGTGGATGGTAACTTATCAACAT 1570
Db 132138 TTGAATTAATTAATGATCTGCCAGCAAGGATGGAATGTGGATGGTAACTTATCAACAT 132079
Qy 1571 ATTTTGACATCAATCTGTTTTTGGATATAATTTTAAAACTGTTTACAAAATCTGGGA 1630
Db 132078 ATTTTGGCATGAATCTGTTTTTGGATATACTTTTAAAACTGTTTAAAAAATCTGGGA 132019
Qy 1631 AGAGGAGAAATAGTGTGTTTTCTTCAATTTGATGAGATATTTGCACAATGGTTCCGCAAAAGC 1690
Db 132018 AGA-GAGAAATAGTGTGTTTTCTTCAATTTGATGAGATATTTGCACAATGGTTTGGCAAAAGC 131960
Qy 1691 AGAACAATATCCGATCTATTTTAACTCAAGGAAATCTGAGATTTATCTGACATCA 1750
Db 131959 AGAACAATATCCGATCTATTTTAACTCAAGGAAATCTGAGATTTATCTGACATCA 131900
Qy 1751 TGGACCTCAGATCTCGGACCAACCCCATTTGCAATGAGCTTTGCAAGTTTCAAAATCTAC 1810
Db 131899 TGGACCTCAGATCTCGGACCAACCCCATTTGCAATGAGCTTTGCAAGTTTCAAAATCTAC 131840
Qy 1811 TGGG-----1814
Db 131839 TAGGATAGGTGGGTGGCGGCTCGCGCTGTAAACCCAGACATTTTGGGAGGCGGAG 131780
Qy 1815 -----1814
Db 131779 AGGGTGGATCACCTGAGGTGAGAAAGTTCAAGATCAGCTTAGTCAACATGGTGAACCCAG 131720
Qy 1815 -----1814
Db 131719 TCTCTATAAATATACAAAATTTAGCCGGGCAACCGTGTGGGCACTGTGTAATCCAGCTA 131660
Qy 1815 -----1814
Db 131659 CTCGCGAGCTGAGGACGAGATCGCTTTGAATCCGGGAGGCGGAGGTTGCGAGTGAGTCG 131600
Qy 1815 -----1814
Db 131599 AGATCACCCCATTTGCTCCGGCTCGGGAACAAAGAGCGAAATCTTCGTCTAAAAAATA 131540
Qy 1815 -----GATAAATCTACATCTGAAGCTTCCTCAGAAACCCA 1851
Db 131539 GAAAGAAAGAAATCTTACTGAGATAAATGACATCTGAAGACTTCTCTCAGAAACCCA 131480

Query Match 59.2%; Score 1897.6; DB 10; Length 3229; Best Local Similarity 82.3%; Pred. No. 0; Matches 2477; Conservative 0; Mismatches 429; Indels 102; Gaps 22;			
QY 55	CGGGTCCTCGCTCGCGCGGGCGGCGGCGAGCG- -GGACCTGGGACTAGCGAACC 112	Db	1030
DB 16	CGAGGCGCAGACTCTCGGCTTAGTCGGCCGCGCGCCGAGGCTCCGGCGCGCGCGGAGCG 75	QY	1127
QY 113	GGAGCAGCAGCATATAAAATAAATCCATCAGAAATGACACCTTCTCAGGTGCTCTTTGAAA 172	Db	1090
DB 76	GAGTGCAGAAATCTTAAATAATCCATCAGAAATGACACCTTCTCAGGTCACTTTGAAA 135	QY	1187
QY 173	TAAGAGGAATCTTTTACAGAGAAAGTTTTTGGGATATGCGAAGTGTGATGCTTTGG 232	Db	1150
DB 136	TAAGAGGAATCTTTTACAGAGAGGCTTTTGCATGTGTGAAAATCTGTGATGCTTTGG 195	QY	1246
QY 233	GAAATCTGGAATCTCAAAATGCTGGCTCTTCTCCAGAGAAATGACACAGGTGAAAGCA 292	Db	1247
DB 196	GAACTGGAGTCTCTAAATGCTGTG- - -CTCTTACTGAGAGTGAGACAGGCGAAG- - 250	QY	1269
QY 293	TGCTATGGAAGCAACCATTTGTACTCAGTGAAGGAGTATCAGTTCAGTATCGCTACTTCA 352	Db	1307
DB 251	-TGATGGAAGCAGTGATTTGTTCTTAGTAGAGGATGTCGTAAGTACCGCTACTTCA 309	QY	1366
QY 353	AAGGTAATCTTTTACAGCAAGATATCGGTGTCATGTCAAGTATGATTTCAAGT 412	Db	1329
DB 310	GAGGCTGCTTTTATAGAACCAAGACTATCGGTGTCATGTCAAGTATGATTTCAAGT 369	QY	1426
QY 413	GGGAGTCACTACACCAAGCATCAATCAACCCCTTTAGAAAGCAAGAAATATATGAGG 472	Db	1389
DB 370	GGGAGTCACTACACCAAGCATCAATCAACCCCTTTAGAAAGCAAGAAATATATGAGG 429	QY	1486
QY 473	ATGGCAATTTTGGAAATCCAAATGTTGTTGAAATCTGGAATCTGGATGCTGCAATGTC 532	Db	1449
DB 430	ATGGCAATTTTGGAAATCCAAATGTTGTTGAAATCTGGAATCTGGATGCTGCAATGTC 489	QY	1546
QY 533	AGACTGAAATAGAGTGGCTGCTGCAATTTCTGAGAAACCTCTGTTTCAAATTAACCA 549	Db	1509
DB 490	AGACTGAAATAGAGTGGCTGCTGCAATTTCTGAGAAACCTCTGTTTCAAATTAACCA 549	QY	1547
QY 593	AAAAATTAATAAGTATAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAG 609	Db	1569
DB 550	AAAAATTAATAAGTATAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAG 609	QY	1666
QY 652	-----GACGATGATAGGTTATCTCCACCTGTAATCTCAGCAAAATGTCCTAGCTGGAG 706	Db	1629
DB 610	ATGACGACATGATAGGTTATCTCCACCTGTAATCTCAGCAAAATGTCCTAGCTGGAG 669	QY	1726
QY 707	TATCCTTAATAAGCGCAATGATTTCAAGTCAGGCAATTCAGCGCGGAGTGTGTTATG 766	Db	1689
DB 670	TATCCTTAATAAGTGAATGATTTCAAGTCAGGCAATTCAGCGCGGAGTGTGTTATG 729	QY	1786
QY 767	GCTTGCAGCTGATCGTTGGAAGAGTATCAGTGAAGTATCAGCAAAATGTCCTAGCTGG 826	Db	1749
DB 730	GCTTGCAGCTGATCGTTGGAAGAGTATCAGTGAAGTATCAGCAAAATGTCCTAGCTGG 789	QY	1846
QY 827	AATCAATCTTGTATTTTTCGAAAGAGTATCAGTGAAGTATCAGCAAAATGTCCTAGCT 886	Db	1809
DB 790	AATCAATCTTGTATTTTTCGAAAGAGTATCAGTGAAGTATCAGCAAAATGTCCTAGCT 849	QY	1906
QY 887	TTCTTGACATGTTGGTATCAGCTTGTCTTATCATCCCAATGCTGAGAGTGAAGAA 946	Db	1869
DB 850	TTCTTGACATGTTGGTATCAGCTTGTCTTATCATCCCAATGCTGAGAGTGAAGAA 909	QY	1966
QY 947	GTCGTGGAATCTTACTCTTCCATCATGAGCGAAATTTCCCGGAAACCAATAGGCAAG 1006	Db	1929
DB 910	GCGCTGGAATCTTACTCTTCCATCATGAGCGAAATTTCCCGGAAACCAATAGGCAAG 969	QY	2026
QY 1007	TCGAGTTCACATATAATTTTAAAGCCATTAACAGGATACAGTTTGTGATGAATCTT 1066	Db	1989
DB 970	TCGAGTTCATTTTATCATCATCAAGCCATTAACAGGATATAGTTTCTTATGCACTT 1029	QY	2086
QY 1067	CATTTTCCAGTATTTGGAAGCCAAAGATACCATTTGGATGTTGGCCATCGAGGTGAGGAA 1126	Db	2049

Db	2110	ATGTGGAAACGCTTAGTTCTCTAGTGCACAGAGGACATTCAGAGGCTCTCCCTCGCGCT- 2168
Qy	2207	CTGGGTATTCATTTTTCATCACGTAGCATTTGTGATCTATGCTCTTTTGGGCTTCTCAGTT 2266
Db	2169	---GAGGTTCCGTCCTCACCACTGAACCGGT-----CGCCTCTTAGGTTTCTCAGTTC 2219
Qy	2267	CAATGAAGCAATATGAAGTATTTAACTCTTTTCACCTACAGTTCTTCGCAAGTATGCTATTT 2326
Db	2220	CAATGAAGCAATATGAAGTATTTTA--CTATCATTTACAGTTCCCGCAAGAATATCAAGT 2277
Qy	2327	AAATTT-----ACTTGGCCAGGTATTAATTTGCCAGTCAGTCTCTTTTATAGTGAGAAAAT 2379
Db	2278	ACATATTTATCACTTGTCCAGGTATAAATTACCAATCAGTCTCTGTACAA-----2327
Qy	2380	TATTTGGTTAGTAATATAAATATTTTAAACTAAATATATAAATCTATAAATGTTTAAACATAT 2439
Db	2328	-----ATGTTAAACACITTTAATGAGAGATCTAAGCCCTATAATGGTCA---AT 2372
Qy	2440	GTTTCATTTAAAGCATAGCACTTTGAAATTAACATATATAAATAGCTCATATTTTACCTTAC 2499
Db	2373	CTTCATTTAAAGCATATACTTTGGAATTAGCTATATAAATATCTCAT-----AGTTCA 2426
Qy	2500	AGCTTTTCATTTGATTCAGGTCGAAATCTTTTACGACCTTAAGGAAATAGCTATGCTAAT 2559
Db	2427	GGCTTTTCATTTGATTAGTCTTAAATCTTCAGTGTCTTGAGAAAATAGAC--TGCATAAT 2484
Qy	2560	TATACCTGACCATGAAAATAATAGTACCTCAATGATGATCATTTGACCTGGTGATTTCCA 2619
Db	2485	TATACCTGACCATGAAAATAATAGTACCTCAAGTGCATGCATTTGCACCTTGGTGGCTCCA 2544
Qy	2620	ACTGCACAAATCTTTTGTGCCATCTTTGTATATAGGTATTTTTCATGGGTTGATGACAC 2679
Db	2545	GCTGCACAAGTCTGTGT-CATCCATGTACATAGGT-GTCTTTACATGGGTTGATAGAAAC 2602
Qy	2680	ACAACACCAATTTTCATTCAGTATGAACCTTGAGGCTGCTGCCATTTTTCACCTTAACCA 2739
Db	2603	A-TGCACCTAGGCTCTCTTTAGTATAAACC-TCAGACTGCTGCCATTTCCCACTGACCCAA 2660
Qy	2740	ACCAGCCTGAAGGTGAACCTCGAACTTGTTTTCATAAATCTTTCAAAAGTTGTTTACAT 2799
Db	2661	ACCAGCCTGCAAGTGAACCTCAAACTGTTTTCATAGCTGTTCAAGATTTTA-----2714
Qy	2800	CAATGTTAAAAATTTCCAAAATGCTGCAGGTAATTTAATGTATAAAATATTAGTAAAGAAA 2859
Db	2715	-----AAGGTTCCAGAAATGCTGCAGGGAACCTTAATGTATAAAGTATTGTGAAG---A 2764
Qy	2860	AGTATGTAATGCATCTTAGTAGAATAGATCACAACTACAATCAATTCAGTGCATGTC 2919
Db	2765	GGTATATATTGCATATATAGTCGTGTAGATCAGAATGTTAAATTTGACTCCGTCGATGC 2824
Qy	2920	TTTAG----GTCTTAAGCATGAGATTGTATACATGTTTACTGTAGTCTCTTCATCTGTGG 2975
Db	2825	TTTAGGTTTGTTTTAAGCATGAGTTGTACATGTTTACT-----GTCTTGCATC--TGG 2877
Qy	2976	TGCTAGGTGAGTATGAGAAGATGTCAAGGACTTGGACGTATTTTGTGTGCTTAAAAAAA 3035
Db	2878	TGCTAGGTGAGT-----GAGATGTTAAGGACTGAAAAATATTTTGTGCGCTAAAAAGAGTA 2932
Qy	3036	GGCTGTTT 3043
Db	2933	TGCTTATT 2940

RESULT 14	BD156897	1803 bp	DNA	linear	PAT 17-JAN-2003			
LOCUS	BD156897	Primer for synthesizing full-length cDNA and use thereof.						
DEFINITION	BD156897							
ACCESSION	BD156897							
VERSION	BD156897.1	GI:27862655						
KEYWORDS	JP 2002191363-A/11740.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							

REFERENCE AUTHORS	TITLE JOURNAL	COMMENT	FEATURES source	ORIGIN
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
1 (bases 1 to 1803)				
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.	Primer for synthesizing full-length cDNA and use thereof	Patent: JP 2002191363-A 11740 09-JUL-2002;		
HELIX RESEARCH INSTITUTE				
OS Homo sapiens (human)				
PN JP 2002191363-A/11740				
PD 09-JUL-2002				
PF 28-JUL-2000 JP 2000280990				
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU				
PI SAITO,				
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,				
PI KEIICHI NAGAI, TETSUJI OTSUKI				
PC				
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/				
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC				
Primer for synthesizing full-length cDNA and use thereof FH Key				
Location/Qualifiers				
FT CDS				
Location/Qualifiers				
1..1803				
/organism="Homo sapiens"				
/mol_type="genomic DNA"				
/db_xref="taxon:9606"				
Query Match 56.2%; Score 1803; DB 6; Length 1803;				
Best Local Similarity 100.0%; Pred. No. 0;				
Matches 1803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY 1379 ATGTGACTGCACCTGAATCTAAGGATCGGAAAGAAATCTGTGGTTCAGGAGGAAATTCCT 1438				
DB 1 ATGTGACTGCACCTGAATCTAAGGATCGGAAAGAAATCTGTGGTTCAGGAGGAAATTCCT 60				
QY 1439 TTTCAGAAATCAGCCATTTCTTCTTAAGATGGTTTAAAGATCTTTAGAGTCTTTGGCAGAGATG 1498				
DB 61 TTTCAGAAATCAGCCATTTCTTCTTAAGATGGTTTAAAGATCTTTAGAGTCTTTGGCAGAGATG 120				
QY 1499 TAGGTTTAAACATTTGAATATAATGGATCTGCCAGCAAGGATGGAAATGGGATGGTA 1558				
DB 121 TAGGTTTAAACATTTGAATATAATGGATCTGCCAGCAAGGATGGAAATGGGATGGTA 180				
QY 1559 ACTTATCAACATATTTGACATGAATCTGTTTGGATATAATTTTAAAACTGTTTATG 1618				
DB 181 ACTTATCAACATATTTGACATGAATCTGTTTGGATATAATTTTAAAACTGTTTATG 240				
QY 1619 AAAATTTCTGGAGAGGAGAAATAGTGTCTTCTCAATTTGATGTCAGATATTTGCACAATGG 1678				
DB 241 AAAATTTCTGGAGAGGAGAAATAGTGTCTTCTCAATTTGATGTCAGATATTTGCACAATGG 300				
QY 1679 TTGGCAAAAGCAGAAACAAATATCCGATACATATTTTAACTCAAGGAAATCTGAGATT 1738				
DB 301 TTGGCAAAAGCAGAAACAAATATCCGATACATATTTTAACTCAAGGAAATCTGAGATT 360				
QY 1739 ATCTGTGAATCATGACCTCAGATCTCGCAACCCCAATGTCATGAGCTTTGACAGT 1798				
DB 361 ATCTGTGAATCATGACCTCAGATCTCGCAACCCCAATGTCATGAGCTTTGACAGT 420				
QY 1799 TTGAAATCTACTGGGGATAAATGTACATCTGGAAGACTTTGTCAGAAACCCATCTTATA 1858				
DB 421 TTGAAATCTACTGGGGATAAATGTACATCTGGAAGACTTTGTCAGAAACCCATCTTATA 480				
QY 1859 TTTCAAGAGGCAAAAGCTAAGGCACTAGTGCATATTTCTGTGGGGTGTATGATCAATGATC 1918				
DB 481 TTTCAAGAGGCAAAAGCTAAGGCACTAGTGCATATTTCTGTGGGGTGTATGATCAATGATC 540				
QY 1919 CTGAAAAACAGAGGAAATTTGAAGAACTTTGAGTGTAAATCGTCTAAATTTATGATAGTAT 1978				
DB 541 CTGAAAAACAGAGGAAATTTGAAGAACTTTGAGTGTAAATCGTCTAAATTTATGATAGTAT 600				

1979 ATGATTTGGATGCTGTAACCAACCAATATATTTCCAGTGGAGCAATTTGGAAACGCTGAAGC 2038
601 ATGATTTGGATGCTGTAACCAACCAATATATTTCCAGTGGAGCAATTTGGAAACGCTGAAGC 660
2039 AGGAATTCAGAGCTTAAGAGCTGTTTGTGTCACCTGTTAGCGCTGTTTTCCTCAT 2098
661 AGGAATTCAGAGCTTAAGAGCTGTTTGTGTCACCTGTTAGCGCTGTTTTCCTCAT 720
2099 CTTTGTGCGGAGTCTGATATCCATGTGGATGTCACCAACGCAATGATAACGTGGAGAATG 2158
721 CTTTGTGCGGAGTCTGATATCCATGTGGATGTCACCAACGCAATGATAACGTGGAGAATG 780
2159 CTTAGTTTTTATGACAGAGTCAATTTTGGGGGCGTGACCGCTGTTCTGGGTATTCAT 2218
781 CTTAGTTTTTATGACAGAGTCAATTTTGGGGGCGTGACCGCTGTTCTGGGTATTCAT 840
2219 TTTTCATCACTAGAGCTGTTGATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAAT 2278
841 TTTTCATCACTAGAGCTGTTGATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAAT 900
2279 AATGAAGTATTTAACTCTTTCACTACAGTCTTTGCAAGTATGCTATTTAAATTTACTTGGC 2338
901 AATGAAGTATTTAACTCTTTCACTACAGTCTTTGCAAGTATGCTATTTAAATTTACTTGGC 960
2339 CAGGTATTAATTCAGCTAGTCTCTTTATAGTGAAGAAATTTATTTGGTGTAGTAATAAA 2398
961 CAGGTATTAATTCAGCTAGTCTCTTTATAGTGAAGAAATTTATTTGGTGTAGTAATAAA 1020
2399 TATTTTAACTAAATATATAATCTATAATGTTAAACATATGTTCAATTAAGCAATAGCA 2458
1021 TATTTTAACTAAATATATAATCTATAATGTTAAACATATGTTCAATTAAGCAATAGCA 1080
2459 CTTTGAATTAATCAATATATAATAGTCTATTTTACATTTACAGTCTTTTCAATTTGATCAGG 2518
1081 CTTTGAATTAATCAATATATAATAGTCTATTTTACATTTACAGTCTTTTCAATTTGATCAGG 1140
2519 TCTGAATCTTTAGCACTTAAGAAATGACATATGCAATTAATTTACCTGACCATGAATAA 2578
1141 TCTGAATCTTTAGCACTTAAGAAATGACATATGCAATTAATTTACCTGACCATGAATAA 1200
2579 AATAAGTACCTCAATGACATGATTTGCACTGCTGATTTCACTGACCAATCTTTGTCG 2638
1201 AATAAGTACCTCAATGACATGATTTGCACTGCTGATTTCACTGACCAATCTTTGTCG 1260
2639 CATCTTGTATATAGGTATTTTATACATGGGTTGACATGACCAACACCAATTTTCATTTCA 2698
1261 CATCTTGTATATAGGTATTTTATACATGGGTTGACATGACCAACACCAATTTTCATTTCA 1320
2699 GTATGAACCTTGAGGCTGCTGCAATTTTTCACATTAACCAACCCAGCTGGAAGTGAACC 2758
1321 GTATGAACCTTGAGGCTGCTGCAATTTTTCACATTAACCAACCCAGCTGGAAGTGAACC 1380
2759 TCGAACTTGTTCATTAATCTTTCAAAAGTGTGTTTACATCAATGTTAAATTTCAAAA 2818
1381 TCGAACTTGTTCATTAATCTTTCAAAAGTGTGTTTACATCAATGTTAAATTTCAAAA 1440
2819 TGTCTGAGGCTAATTTAATGTTAATAATATAGTAAAGAAATGATGATTTGCATCTTA 2878
1441 TGTCTGAGGCTAATTTAATGTTAATAATATAGTAAAGAAATGATGATTTGCATCTTA 1500
2879 GPAGAATAGATCAACATACAAATTTCAATTCAGTGCATGCTTTTAGGCTTAAGCATGAG 2938
1501 GPAGAATAGATCAACATACAAATTTCAATTCAGTGCATGCTTTTAGGCTTAAGCATGAG 1560
2939 ATTGTACATGTTTACGTTAGTCTTGTGATCTGCTGCTAGGTGAGTATGAGAAGATG 2998
1561 ATTGTACATGTTTACGTTAGTCTTGTGATCTGCTGCTAGGTGAGTATGAGAAGATG 1620
2999 TCAAGGACTGGAAGTATTTTGTGCTCAAAAAAAGGCTGTTTGTAGCGCTTTTAAAT 3058
1621 TCAAGGACTGGAAGTATTTTGTGCTCAAAAAAAGGCTGTTTGTAGCGCTTTTAAAT 1680
3059 ATGCTTATTTTGTGCTCTCACTACCTATTTACACACTGTTTGTGCTTTGTTGTTT 3118

1681 ATGCTTATTTTGTGCTCTCACTACCTATTTACACACTGTTGCTTTGCGGTTTGTTTT 1740
3119 GTATGTCGCTGTTTATATACAGTAGTTAAATTTCCATGACAGAAAAATAAATGTCCTGAAT 3178
1741 GTATGTCGCTGTTTATATACAGTAGTTAAATTTCCATGACAGAAAAATAAATGTCCTGAAT 1800
3179 CTC 3181
1801 CTC 1803
RESULT 15
AX877873 1803 bp DNA linear PAT 17-DEC-2003
Sequence 12778 from Patent EP1074617.
DEFINITION
ACCESSION AX877873
VERSION AX877873.1 GI:40032609
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 12778 07-FEB-2001;
RESEARCH Association for Biotechnology (JP)
FEATURES
Location/Qualifiers
1..1803
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
93..785
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89913.1"
/db_xref="GI:40032610"
/translation="MWLESLPEDVGFNIEIKWIQOORDGMWDLNLSYFDMNLFDDII
LKTVLNSGKRIIVPSFDADICTVRQKQKYPILPLTQKSELYPELMDLRSRTTP
IAMSFAOPENLLGINVHTEDLLRNPSTIQEAKAGLVIFCHGDDTNDPDRKLELGS
VNGLYDRIYDMPQPNFQVLEQLUKQELPELKSCLCTPTVSRFVSSLCGESDIH
VDANGIDNVENA"
ORIGIN
Query Match 56.2%; Score 1803; DB 6; Length 1803;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1379 ATGTGACTGCACCTGAAATCTAAGGATCGAAAGAAATCTGTGTTCCAGGAGGAAAAATTCCT 1438
DB 1 ATGTGACTGCACCTGAAATCTAAGGATCGAAAGAAATCTGTGTTCCAGGAGGAAAAATTCCT 60
QY 1439 TTTTCAGAAATCAGCCATTTCTTCTCTTAAGATGGTTTTAGAGTCTTTTGCAGAGATG 1498
DB 61 TTTTCAGAAATCAGCCATTTCTTCTCTTAAGATGGTTTTAGAGTCTTTTGCAGAGATG 120
QY 1499 TAGGGTTTAACTTGAATTAATTAAGATCTGCCAGCAAGGATGGAATGTTGGATGTA 1558
DB 121 TAGGGTTTAACTTGAATTAATTAAGATCTGCCAGCAAGGATGGAATGTTGGATGTA 180
QY 1559 ACTTATCAACATATTTTGCATGATCTGTTTGTGATATATTTTAAAAAACTGTTTGTAG 1618
DB 181 ACTTATCAACATATTTTGCATGATCTGTTTGTGATATATTTTAAAAAACTGTTTGTAG 240
QY 1619 AAAATTTCTGGGAAGGAGAGATAGTGTCTTCTCATTTGATGACAGATATTTGCACAATGG 1678
DB 241 AAAATTTCTGGGAAGGAGAGATAGTGTCTTCTCATTTGATGACAGATATTTGCACAATGG 300
QY 1679 TTCGCAAAAGCAGAACAAATATCCGATACCTATTTTAACTCAAGGAAAAATCTGAGATTT 1738
DB 301 TTCGCAAAAGCAGAACAAATATCCGATACCTATTTTAACTCAAGGAAAAATCTGAGATTT 360

QY 1739 ATCTGAACTCATGGACCTCAGATCTCGGACAAACCCCATTTGCAAGTGGACAGT 1798
DB 361 ATCTGAACTCATGGACCTCAGATCTCGGACAAACCCCATTTGCAAGTGGACAGT 420
QY 1799 TTGAAATCTACTCGGGATAAATGTACATCTGAGACTTGTCTCAGAAACCCATCTTATA 1858
DB 421 TTGAAATCTACTCGGGATAAATGTACATCTGAGACTTGTCTCAGAAACCCATCTTATA 480
QY 1859 TTCAAGAGGCAAAAGCTTAAGGACTAGTCTATTTCTGCTGGGGTGATGATACCAATGATC 1918
DB 481 TTCAAGAGGCAAAAGCTTAAGGACTAGTCTATTTCTGCTGGGGTGATGATACCAATGATC 540
QY 1919 CTGAAAACAGAAAGAAATTTGAAGAACTTGGAGTTAATGTCTAATTTATGATAGATAT 1978
DB 541 CTGAAAACAGAAAGAAATTTGAAGAACTTGGAGTTAATGTCTAATTTATGATAGATAT 600
QY 1979 ATGATTGGATGCCCTGAACAAACCAATATATTTCCAAGTGGAGCAATTTGGAACGCCCTGAAGC 2038
DB 601 ATGATTGGATGCCCTGAACAAACCAATATATTTCCAAGTGGAGCAATTTGGAACGCCCTGAAGC 660
QY 2039 AGGAATGCCAGAGCTTAAGAGCTGTTGTGCTCCACTGTTAGCGGCTTTGTTCCTCAT 2098
DB 661 AGGAATGCCAGAGCTTAAGAGCTGTTGTGCTCCACTGTTAGCGGCTTTGTTCCTCAT 720
QY 2099 CTTTGTGGGGAGTCTGATATCCATGTGGATGCCAAGCGCATTTGATTAACGTGGAGAAATG 2158
DB 721 CTTTGTGGGGAGTCTGATATCCATGTGGATGCCAAGCGCATTTGATTAACGTGGAGAAATG 780
QY 2159 CTTAGTTTTTATGCAAGAGTCAATTTGGGGGGTGCACCGCTGTTCTGGGTATTCAT 2218
DB 781 CTTAGTTTTTATGCAAGAGTCAATTTGGGGGGTGCACCGCTGTTCTGGGTATTCAT 840
QY 2219 TTTTCACTACCTGAGCATTTGTGATCTATGCCCTTTTGGGGCTTCTCAGTTCAATGAAGCAAT 2278
DB 841 TTTTCACTACCTGAGCATTTGTGATCTATGCCCTTTTGGGGCTTCTCAGTTCAATGAAGCAAT 900
QY 2279 AATGAAGTATTTAACTCTTTTCACTACAGTTCCTGCAAGTATGCTATTTAAATTAATTCTGGC 2338
DB 901 AATGAAGTATTTAACTCTTTTCACTACAGTTCCTGCAAGTATGCTATTTAAATTAATTCTGGC 960
QY 2339 CAGGTATAATGCCAGTCACTCTTTTATAGTGAGAAATTTATTTGGTTAGTATATAAA 2398
DB 961 CAGGTATAATGCCAGTCACTCTTTTATAGTGAGAAATTTATTTGGTTAGTATATAAA 1020
QY 2399 TATTTTAACTAAATATATAAATCTATAATGTTAAACATATGTTTAAAGCATAGCA 2458
DB 1021 TATTTTAACTAAATATATAAATCTATAAATGTTTAAACATATGTTTAAAGCATAGCA 1080
QY 2459 CTTTGAATTAATTAATAAATAGCTCATATTTTACACTTTACAGCTTTTCATTTGATCAGG 2518
DB 1081 CTTTGAATTAATTAATAAATAGCTCATATTTTACACTTTACAGCTTTTCATTTGATCAGG 1140
QY 2519 TCTGAAATCTTTAGCACTTAAGAAATGACATATGCAATTAATTAACCTGACCATGAAAA 2578
DB 1141 TCTGAAATCTTTAGCACTTAAGAAATGACATATGCAATTAATTAACCTGACCATGAAAA 1200
QY 2579 AATAAGTACCTCAATGCAATGCAATTTGCACTGGTATCCAACTGCACAAATCTTTGTGC 2638
DB 1201 AATAAGTACCTCAATGCAATTTGCACTGGTATCCAACTGGTATCCAACTGCACAAATCTTTGTGC 1260
QY 2639 CATCTTTGATATAGGTATTTTATCATGGTTGACATGACACAAACCAATTTTCATTCA 2698
DB 1261 CATCTTTGATATAGGTATTTTATCATGGTTGACATGACACAAACCAATTTTCATTCA 1320
QY 2699 GTATGAACCTTTGAGGCTGCTGCAATTTTCCACTTAACCAACCGCCCTGAAGTGAACC 2758
DB 1321 GTATGAACCTTTGAGGCTGCTGCAATTTTCCACTTAACCAACCGCCCTGAAGTGAACC 1380
QY 2759 TCGAACTTTGTTTCATAAATCTTTCAAAAGTCTTTTACATCAATGTTTAAATTTTCAAAA 2818
DB 1381 TCGAACTTTGTTTCATAAATCTTTTCAAAAGTCTTTTACATCAATGTTTAAATTTTCAAAA 1440

QY 2819 TGCTGCAGGGTAATTTAATGTATATAAATATTTAGTAAGAAAAAGTATGTATTGCTACTTA 2878
DB 1441 TGCTGCAGGGTAATTTAATGTATATAAATATTTAGTAAGAAAAAGTATGTATTGCTACTTA 1500
QY 2879 GTAGAATAGATCAAAACATACAAATTTCAATTCAGTGCATGCTTTAGGTGTTAAGCATGAG 2938
DB 1501 GTAGAATAGATCAAAACATACAAATTTCAATTCAGTGCATGCTTTAGGTGTTAAGCATGAG 1560
QY 2939 ATTGTACATGTTTACTGTTAGTCTCTGCTAGTGTGCTAGGTGATATGAGAAGATG 2998
DB 1561 ATTGTACATGTTTACTGTTAGTCTCTGCTAGTGTGCTAGGTGATATGAGAAGATG 1620
QY 2999 TCAAGGACTTGGACGTATTTTGTGCTAAAAAAGGCTGTTTGTAGGCGTTTAAAAAT 3058
DB 1621 TCAAGGACTTGGACGTATTTTGTGCTAAAAAAGGCTGTTTGTAGGCGTTTAAAAAT 1680
QY 3059 ATGCTTATTTTGTGCTGCTCTCACCTATTTACACACTGTTGCTTCTGCTGGTGTGTTTT 3118
DB 1681 ATGCTTATTTTGTGCTGCTCTCACCTATTTACACACTGTTGCTTCTGCTGGTGTGTTTT 1740
QY 3119 GTATGTGCGTGTATATACAGTAGTTAAATTTCCATGCAAAAAATAAATGTCCTGAATTT 3178
DB 1741 GTATGTGCGTGTATATACAGTAGTTAAATTTCCATGCAAAAAATAAATGTCCTGAATTT 1800
QY 3179 CTC 3181
DB 1801 CTC 1803

Search completed: July 2, 2005, 16:56:09
Job time : 13678 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2005, 16:56:17 ; Search time 1861 Seconds
(without alignments)
10805.278 Million cell updates/sec

Title: US-10-047-855-4

Perfect score: 3206

Sequence: 1 gtcacccacgcgtccggc.....aaaaaaaaagcgccgcgc 3206

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3206	100.0	3206	16	US-10-047-855-4
2	3206	100.0	3206	17	US-10-426-776-34
3	3178	99.1	3265	21	US-10-887-553A-1007
4	3166	98.8	3499	10	US-09-814-353-21000
5	3164.4	98.7	3499	20	US-10-357-930-21700
6	3164.4	98.7	3499	20	US-10-357-930-29632
7	2481.8	77.4	2738	17	US-10-426-776-32

8	2019	63.0	2019	21	US-10-772-636-25	Sequence 25, Appl
9	1471.2	45.9	3381	17	US-10-426-776-39	Sequence 39, Appl
10	991	30.9	2393	17	US-10-426-776-48	Sequence 48, Appl
11	522.4	16.3	524	9	US-09-796-692-4214	Sequence 4214, Ap
12	522.4	16.3	524	14	US-10-040-862-4214	Sequence 4214, Ap
13	522.4	16.3	524	17	US-10-057-475B-4214	Sequence 4214, Ap
14	522.4	16.3	524	17	US-10-154-884B-4214	Sequence 4214, Ap
15	522.4	16.3	524	19	US-10-764-324-4214	Sequence 4214, Ap
16	486.4	15.2	546	9	US-09-777-564-71	Sequence 71, Appl
17	486.4	15.2	546	14	US-10-015-219-71	Sequence 71, Appl
18	469.4	14.6	471	9	US-09-998-598-1849	Sequence 1849, Ap
19	427	13.3	427	10	US-09-918-995-6192	Sequence 1692, Ap
20	414.8	12.9	520	10	US-09-814-353-16669	Sequence 16669, A
21	413.8	12.9	436	9	US-09-796-692-8067	Sequence 8067, Ap
22	413.8	12.9	436	14	US-10-040-862-8067	Sequence 8067, Ap
23	413.8	12.9	436	17	US-10-057-475B-8067	Sequence 8067, Ap
24	413.8	12.9	436	17	US-10-154-884B-8067	Sequence 8067, Ap
25	413.8	12.9	436	19	US-10-764-324-8067	Sequence 8067, Ap
26	365.6	11.4	403	9	US-09-777-564-183	Sequence 183, App
27	365.6	11.4	403	14	US-10-015-219-183	Sequence 183, App
28	363.4	11.3	396	10	US-09-814-353-3977	Sequence 3977, Ap
29	363.4	11.3	396	10	US-09-814-353-10285	Sequence 10285, A
30	326.6	10.2	425	20	US-10-357-930-35883	Sequence 35883, A
31	326.6	10.2	425	20	US-10-357-930-44946	Sequence 44946, A
32	326	10.2	327	20	US-10-357-930-15054	Sequence 15054, A
33	296.4	9.2	478	9	US-09-867-701-2369	Sequence 4363, Ap
34	291	9.1	510	9	US-09-867-701-2369	Sequence 2369, Ap
35	271	8.5	359	20	US-10-357-930-5885	Sequence 5885, Ap
36	208.4	6.5	463	9	US-09-864-761-2380	Sequence 2380, Ap
37	193	6.0	193	9	US-09-864-761-31997	Sequence 31997, A
38	185.8	5.8	1121	14	US-10-198-846-11446	Sequence 11446, A
39	166	5.2	401	9	US-09-864-761-15483	Sequence 15483, A
40	163.4	5.1	323	9	US-09-728-445-622	Sequence 622, App
41	151	4.7	151	9	US-09-864-761-19114	Sequence 19114, A
42	138	4.3	429	9	US-09-864-761-15673	Sequence 15673, A
43	122	3.8	122	9	US-09-864-761-32178	Sequence 32178, A
44	120.2	3.7	354	17	US-10-242-535A-35647	Sequence 35647, A
45	120.2	3.7	354	18	US-10-085-783A-35647	Sequence 35647, A

ALIGNMENTS

RESULT 1
US-10-047-855-4
; Sequence 4, Application US/10047855
; Publication No. US20030165863A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Lillian Wei-Ming
; TITLE OF INVENTION: NARC10 and NARC16, Programmed Cell
; DEATH-ASSOCIATED MOLECULES AND USES THEREOF
; FILE REFERENCE: 35800/242056
; CURRENT APPLICATION NUMBER: US/10/047,855
; PRIOR FILING DATE: 2002-01-15
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; TYPE: DNA
; LENGTH: 3206
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)...(2163)
; NAME/KEY: misc feature
; LOCATION: (1)...(17)
; OTHER INFORMATION: Vector sequence
US-10-047-855-4

Query Match 100.0%; Score 3206; DB 16; Length 3206;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACACCAACGCGTCCGGGCGAGGCAACGGACGGCGGGCGCCGGTACCTCTGCCCGGGT 60
Db 1 GTGACACCAACGCGTCCGGGCGAGGCAACGGACGGCGGGCGCCGGTACCTCTGCCCGGGT 60
QY 61 CCTCGCTCTCGGGGGGGGGCGCGGAGCGAGCGGACCTCGGAGCTAGCGAAACCCGGAGCAG 120
Db 61 CCTCGCTCTCGGGGGGGGGCGCGGAGCGAGCGGACCTCGGAGCTAGCGAAACCCGGAGCAG 120
QY 121 ACATCATAAATAAATCCATCAGAAATGACACCTTCTCAGGTTGGCTTTGAAATAAGAGGA 180
Db 121 ACATCATAAATAAATCCATCAGAAATGACACCTTCTCAGGTTGGCTTTGAAATAAGAGGA 180
QY 181 ACTCTTTTACCAGGAGAAAGTTTTTGGGATATGTGGAAGCTGTGATGCTTTGGGAAACTGG 240
Db 181 ACTCTTTTACCAGGAGAAAGTTTTTGGGATATGTGGAAGCTGTGATGCTTTGGGAAACTGG 240
QY 241 AATCCTCAAAATGTGTGGCTCTTCTTCCAGAGAAATGACACAGGTGAAAGCTCTATGG 300
Db 241 AATCCTCAAAATGTGTGGCTCTTCTTCCAGAGAAATGACACAGGTGAAAGCTCTATGG 300
QY 301 AAAGCAACCATGTACTCAGTAGAGGAGTATCAGTTCAAGTATCGTACTTCAAAGGGTAC 360
Db 301 AAAGCAACCATGTACTCAGTAGAGGAGTATCAGTTCAAGTATCGTACTTCAAAGGGTAC 360
QY 361 TTTTGTAGAACCAAGACTATCGGTGGTCCATGTCTCAAGTGTAGTTCACAAAGTGGAGACT 420
Db 361 TTTTGTAGAACCAAGACTATCGGTGGTCCATGTCTCAAGTGTAGTTCACAAAGTGGAGACT 420
QY 421 CATCTACACCAACGATCAATAACCCCTTTAGAAAGCGAAATATTTATTTGACGATGGACAA 480
Db 421 CATCTACACCAACGATCAATAACCCCTTTAGAAAGCGAAATATTTATTTGACGATGGACAA 480
QY 481 TTTGGAAATCCAAATGGTGTGAAACTCTGGAATCTGGATCTGGATGCTGACATGTCAGACTGAA 540
Db 481 TTTGGAAATCCAAATGGTGTGAAACTCTGGAATCTGGATCTGGATGCTGACATGTCAGACTGAA 540
QY 541 ATAAGATTACGTTTGTGATTAATCTGAAATAACCTCTGTGTCAATAACCAAGAAAAATTA 600
Db 541 ATAAGATTACGTTTGTGATTAATCTGAAATAACCTCTGTGTCAATAACCAAGAAAAATTA 600
QY 601 AAAAAATCTAGATTATAGGTGAAGCTGACACATAGAAAGCGCTGGAGGAAGATGACGATGAT 660
Db 601 AAAAAATCTAGATTATAGGTGAAGCTGACACATAGAAAGCGCTGGAGGAAGATGACGATGAT 660
QY 661 AGGGTATCTCCCACTGACTCCACAAAATGTCCAATAGCTTGGAGATATCCTTAATAAGC 720
Db 661 AGGGTATCTCCCACTGACTCCACAAAATGTCCAATAGCTTGGAGATATCCTTAATAAGC 720
QY 721 GACAAATGAGTTCAAAGTGACGGCAATTCACAGCCGGAGTGTGGTTATGGCTTGCAGCCTGAT 780
Db 721 GACAAATGAGTTCAAAGTGACGGCAATTCACAGCCGGAGTGTGGTTATGGCTTGCAGCCTGAT 780
QY 781 CGTTGGACAGAGTACAGCATACAGAGATGGAACCAAGATAA CCTGGAACATACTTTTGTAT 840
Db 781 CGTTGGACAGAGTACAGCATACAGAGATGGAACCAAGATAA CCTGGAACATACTTTTGTAT 840
QY 841 TTTTTCGAAAGAGATCTCAGTGAGCAGCTAGTTTCAGGGTGATGCCCTTCTCGACATATG 900
Db 841 TTTTTCGAAAGAGATCTCAGTGAGCAGCTAGTTTCAGGGTGATGCCCTTCTCGACATATG 900
QY 901 GGTACAGCTTGTCTCTTATCATCCACCATTTGCTGAGAGTGGAAAGAGTCTGGAAATTCCT 960
Db 901 GGTACAGCTTGTCTCTTATCATCCACCATTTGCTGAGAGTGGAAAGAGTCTGGAAATTCCT 960
QY 961 ACTCTTCCCATATGAGAGAAATTCGGGAAACCAATAGGCGAAAGTGTGACTAT 1020
Db 961 ACTCTTCCCATATGAGAGAAATTCGGGAAACCAATAGGCGAAAGTGTGACTAT 1020
QY 1021 ATAAATTTAAGCAATTTACCAGAGATACAGTTGTGACATGAATCTTCAATTTCCAGAGTAT 1080
Db 1021 ATAAATTTAAGCAATTTACCAGAGATACAGTTGTGACATGAATCTTCAATTTCCAGAGTAT 1080

QY 1081 TGGAAAGCCAAAGTAATACCATTTGGATGTTGGCCATCGAGGTGCAGGAAAATCTTACAACT 1140
Db 1081 TGGAAAGCCAAAGTAATACCATTTGGATGTTGGCCATCGAGGTGCAGGAAAATCTTACAACT 1140
QY 1141 GCCCAGCTGGCTAAAGTTCAAGAAAATACTATTTGCTCTTTTAAAGAAATGCTGCTAGTCA 1200
Db 1141 GCCCAGCTGGCTAAAGTTCAAGAAAATACTATTTGCTCTTTTAAAGAAATGCTGCTAGTCA 1200
QY 1201 GGTGCAGCCCTTTGTAGAATTTGACGTACACCTTTCAAAGGACTTTTGTGCCGTGTATAT 1260
Db 1201 GGTGCAGCCCTTTGTAGAATTTGACGTACACCTTTCAAAGGACTTTTGTGCCGTGTATAT 1260
QY 1261 CATGATCTTACCTGTTGTTGACTATGAAAAGAAAATTTGATGCTGATCCAGTTGAAATTA 1320
Db 1261 CATGATCTTACCTGTTGTTGACTATGAAAAGAAAATTTGATGCTGATCCAGTTGAAATTA 1320
QY 1321 TTTGAAATTTCCAGTAAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCAT 1380
Db 1321 TTTGAAATTTCCAGTAAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCAT 1380
QY 1381 GTGACTGCACTGAAATCTTAAGGATCGGAAAGAAATCTGTGGTTCAGAGGAAAATTCCTTT 1440
Db 1381 GTGACTGCACTGAAATCTTAAGGATCGGAAAGAAATCTGTGGTTCAGAGGAAAATTCCTTT 1440
QY 1441 TCAGAAAAATCAGCCATTTCTTCTTAAGATGTTTGTAGAGTCTTTGCGAGAAGATGTA 1500
Db 1441 TCAGAAAAATCAGCCATTTCTTCTTAAGATGTTTGTAGAGTCTTTGCGAGAAGATGTA 1500
QY 1501 GGGTTTTAAATTTGCAATGAAATGGAATCTGCCAGCAAGGGATGGAATGTGGATGGTAAC 1560
Db 1501 GGGTTTTAAATTTGCAATGAAATGGAATCTGCCAGCAAGGGATGGAATGTGGATGGTAAC 1560
QY 1561 TTATCAACATATTTTGCACATGAAATCTGTTTTTGGATATAATTTTAAATACTGTTTTAGAA 1620
Db 1561 TTATCAACATATATTTTGCACATGAAATCTGTTTTTGGATATAATTTTAAATACTGTTTTAGAA 1620
QY 1621 AATTTCTGGGAAGAGGAGAAATAGTGTCTTCTTAAAGATGTTTGTAGAGTCTTTGCGAGAAGATGTA 1680
Db 1621 AATTTCTGGGAAGAGGAGAAATAGTGTCTTCTTAAAGATGTTTGTAGAGTCTTTGCGAGAAGATGTA 1680
QY 1681 CGGCAAAAGCAGAAACAAATATCCGATACATATTTTAACTCAAGGAAAATCTGAGATTTAT 1740
Db 1681 CGGCAAAAGCAGAAACAAATATCCGATACATATTTTAACTCAAGGAAAATCTGAGATTTAT 1740
QY 1741 CTTGAACTCATGACCTCAGATCTCGGCAACCCCATTTGCAATGAGCTTTGCGACAGTTT 1800
Db 1741 CTTGAACTCATGACCTCAGATCTCGGCAACCCCATTTGCAATGAGCTTTGCGACAGTTT 1800
QY 1801 GAAAACTCTCTGGGGATAAATGTACATCTGAAAGACTTGTCTCAGAAAACCCATCCTATAT 1860
Db 1801 GAAAACTCTCTGGGGATAAATGTACATCTGAAAGACTTGTCTCAGAAAACCCATCCTATAT 1860
QY 1861 CAAGAGCAAAAAGCTAAGGGACTAGTCAATTTCTGTGGGGTGTATGATACCAATGATCCT 1920
Db 1861 CAAGAGCAAAAAGCTAAGGGACTAGTCAATTTCTGTGGGGTGTATGATACCAATGATCCT 1920
QY 1921 GAAAAAGAGAGAAATTTGAAGGAACTTTGAGTTAAATGTTTAAATTTATGATAGATATAT 1980
Db 1921 GAAAAAGAGAGAAATTTGAAGGAACTTTGAGTTAAATGTTTAAATTTATGATAGATATAT 1980
QY 1981 GATTTGGATGCTGAAACCAACCAATAATATTTCCAAAGTGAGCAATTTGGAACCCCTGGAAGCAG 2040
Db 1981 GATTTGGATGCTGAAACCAACCAATAATATTTCCAAAGTGAGCAATTTGGAACCCCTGGAAGCAG 2040
QY 2041 GAAATGGCAGAGCTTAAAGAGCTGTTGTGCCCATCTGTTAGCCGCTTTGTTCCCTCATCT 2100
Db 2041 GAAATGGCAGAGCTTAAAGAGCTGTTGTGCCCATCTGTTAGCCGCTTTGTTCCCTCATCT 2100
QY 2101 TTGTGTGGGGAGTCTGATATCCATGTCGATGCCAACCGCATTTGATAACGTGGAGAAATGCT 2160
Db 2101 TTGTGTGGGGAGTCTGATATCCATGTCGATGCCAACCGCATTTGATAACGTGGAGAAATGCT 2160
QY 2161 TAGTTTTTTATTTGACACAGAGGTCAATTTTGGGGGGTGCACCGCTGTTCTGGGTATTCATTT 2220

Db	2161		TAGT	TTTTATTTG	CGACAGAGGT	CA	TTTTGGGGGGCGGTG	CACCGCTG	TTCTGGGTAT	TTCA	TTT	2222
Qy	2221	TTCA	TCACTGAGCA	TTGTTGAT	CTATGCT	TTTGGGCT	TTCTCAGT	TTCAATGAAGCA	TAATA			2280
Db	2221	TTCA	TCACTGAGCA	TTGTTGAT	CTATGCT	TTTGGGCT	TTCTCAGT	TTCAATGAAGCA	TAATA			2280
Qy	2281	TGAAGTA	TTAACTCTCTT	CAC	TACAGT	TTCTGAAAGT	ATGCTATTTAAAT	TAATCTTTGGCCCA				2340
Db	2281	TGAAGTA	TTAACTCTCTT	CAC	TACAGT	TTCTGAAAGT	ATGCTATTTAAAT	TAATCTTTGGCCCA				2340
Qy	2341	GGTA	TAAATGCGCAGT	CTCTTTA	TAGTCAGAGAAAT	TTTATTTGGT	TAGTAAATATAA	TA				2400
Db	2341	GGTA	TAAATGCGCAGT	CTCTTTA	TAGTCAGAGAAAT	TTTATTTGGT	TAGTAAATATAA	TA				2400
Qy	2401	TTTT	TAAACTAAAT	TATAAAAT	CTATAAT	TGTTTAAACAT	ATGTTCTATTAAAA	AGCATAGC	ACT			2460
Db	2401	TTTT	TAAACTAAAT	TATAAAAT	CTATAAT	TGTTTAAACAT	ATGTTCTATTAAAA	AGCATAGC	ACT			2460
Qy	2461	TTGAAA	TTAACTATATAA	ATAGCTCA	TATTTTAC	CTTACAGCT	TTTTTCATTTGAT	CAGGTC				2520
Db	2461	TTGAAA	TTAACTATATAA	ATAGCTCA	TATTTTAC	CTTACAGCT	TTTTTCATTTGAT	CAGGTC				2520
Qy	2521	TGAAT	CTTTAGCACTTA	AGGAAAT	TGCACTAT	GCATATTA	TATACCTG	ACCATGAA	AAAAA			2580
Db	2521	TGAAT	CTTTAGCACTTA	AGGAAAT	TGCACTAT	GCATATTA	TATACCTG	ACCATGAA	AAAAA			2580
Qy	2581	TAAGT	ACTC	CAAA	TGCACTG	CAATTTG	CCACTGGT	GAATCCAA	CTGCACACAAAT	CTTTGTG	CCCA	2640
Db	2581	TAAGT	ACTC	CAAA	TGCACTG	CAATTTG	CCACTGGT	GAATCCAA	CTGCACACAAAT	CTTTGTG	CCCA	2640
Qy	2641	TC	TTGTATATAG	GTATTTTTT	TACATGGG	TTGACATG	CA	CA	CA	CA	CA	2700
Db	2641	TC	TTGTATATAG	GTATTTTTT	TACATGGG	TTGACATG	CA	CA	CA	CA	CA	2700
Qy	2701	ATGA	ACCTTGAGG	CTGTCGCAT	TTTTTCCATTA	ACCAAC	CCAGCCTGA	AGGTGA	CACTC			2760
Db	2701	ATGA	ACCTTGAGG	CTGTCGCAT	TTTTTCCATTA	ACCAAC	CCAGCCTGA	AGGTGA	CACTC			2760
Qy	2761	GAA	CTTGT	TTTTCATA	AAATCTTTT	CAAAAGT	TGTTTTTAC	ATCAATGTT	AAAAATTTT	C	AA	2820
Db	2761	GAA	CTTGT	TTTTCATA	AAATCTTTT	CAAAAGT	TGTTTTTAC	ATCAATGTT	AAAAATTTT	C	AA	2820
Qy	2821	CTG	CAGGGTA	ATTTAA	TGTATAAAAT	TATAGTA	AGAAAAAGT	ATGAT	TG	CATCTT	AGT	2880
Db	2821	CTG	CAGGGTA	ATTTAA	TGTATAAAAT	TATAGTA	AGAAAAAGT	ATGAT	TG	CATCTT	AGT	2880
Qy	2881	AGAA	TAGAT	CACACAT	CAAAATTTCA	ATTAGTCAT	GCATGCTTTAGG	TGTTAAGC	ATGAT			2940
Db	2881	AGAA	TAGAT	CACACAT	CAAAATTTCA	ATTAGTCAT	GCATGCTTTAGG	TGTTAAGC	ATGAT			2940
Qy	2941	TGT	ACATGTTTAC	TGTTAGGTC	CTTGCACTG	TGGTGTAGGTC	AGTATGAGT	ATGAGAGAT	GTG			3000
Db	2941	TGT	ACATGTTTAC	TGTTAGGTC	CTTGCACTG	TGGTGTAGGTC	AGTATGAGT	ATGAGAGAT	GTG			3000
Qy	3001	AAG	CACTGGA	CGTATTTT	TGTCCTTAAAA	AAAAAAGG	CTGTTTGTAG	CGGTTTTAA	ATAT			3060
Db	3001	AAG	CACTGGA	CGTATTTT	TGTCCTTAAAA	AAAAAAGG	CTGTTTGTAG	CGGTTTTAA	ATAT			3060
Qy	3061	GCT	TATTTTGTG	TGCTCTCA	CTACCTATTA	CACATGTTG	CTTTGTGGG	TTTTGTTTGT				3120
Db	3061	GCT	TATTTTGTG	TGCTCTCA	CTACCTATTA	CACATGTTG	CTTTGTGGG	TTTTGTTTGT				3120
Qy	3121	ATG	TGCGGTG	TGTTATAC	AGTATGTTTAA	ATTTCCATG	TCAGAAAAAT	AAAAATG	TCCTGAA	TTCT		3180
Db	3121	ATG	TGCGGTG	TGTTATAC	AGTATGTTTAA	ATTTCCATG	TCAGAAAAAT	AAAAATG	TCCTGAA	TTCT		3180
Qy	3181	CA	AAAAA	AAAAA	AAAAAGG	CGCGC		3206				
Db	3181	CA	AAAAA	AAAAA	AAAAAGG	CGCGC		3206				

RESULT 2

```

US-10-426-776-34
; Sequence 34, Application US/10426776
; Publication No. US2004009553A1
; GENERAL INFORMATION:
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Williamson, Mark J.
; APPLICANT: Tsia, Fong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Chiang, Lillian Wei-Ming
; APPLICANT: Hunter, John Joseph
; APPLICANT: Wood, Andrew
; APPLICANT: Jenkins, Lorayne P.
; TITLE OF INVENTION: NOVEL 27411, 23413, 22438, 23553,
; TITLE OF INVENTION: 25278, 26212, NARC SC1, NARC 10A, NARC 1, NARC 12, NARC 13,
; TITLE OF INVENTION: NARC17, NARC 25, NARC 3, NARC 4, NARC 7, NARC 8, NARC 11,
; TITLE OF INVENTION: NARC14A, NARC 15, NARC 16, NARC 19, NARC 20, NARC 26, NARC
; TITLE OF INVENTION: 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C, NARC
; TITLE OF INVENTION: 8B, NARC 9, NARC2A, NARC 16B, NARC 1C, NARC 1A, NARC 25,
; TITLE OF INVENTION: 86604 AND 32222 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MP103-0620NIM
; CURRENT APPLICATION NUMBER: US/10/426,776
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 10/229,662
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/795,691
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,517
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 10/105,992
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/406,045
; PRIOR FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 10/314,881
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 09/773,426
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 09/495,823
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/692,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/161,188
; PRIOR FILING DATE: 1999-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 3206
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-10-426-776-14

```


Qy	241	AACTCCTCAA	AATGCTCTGCTGGCTCTTCTTCCAGAGAAATGACACAGGTGAAGCATGCTATGG	300
Db	241			
Qy	241	AACTCCTCAA	AATGCTCTGCTGGCTCTTCTTCCAGAGAAATGACACAGGTGAAGCATGCTATGG	300
Db	241			
Qy	301	AAAGCAACCA	ATTGTACTCAGTGTAGGAGGATATCAGTTCAGTATCGCTACTTCAAAAGGGTAC	360
Db	301			
Qy	301	AAAGCAACCA	ATTGTACTCAGTGTAGGAGGATATCAGTTCAGTATCGCTACTTCAAAAGGGTAC	360
Db	301			
Qy	361	TTTTTTAGAA	CCAAAGACTATTCGGTGGTCCATGTCTCAAGTGAATAGTTTCAAGTGGGAGACT	420
Db	361			
Qy	361	TTTTTTAGAA	CCAAAGACTATTCGGTGGTCCATGTCTCAAGTGAATAGTTTCAAGTGGGAGACT	420
Db	361			
Qy	421	CATCTCAAA	CCACGATCAATFAACCCCTTTAGAAAGCGAAATTTATTATTGACGATGGACAA	480
Db	421			
Qy	421	CATCTCAAA	CCACGATCAATFAACCCCTTTAGAAAGCGAAATTTATTATTGACGATGGACAA	480
Db	421			
Qy	481	TTTTGGAA	TCCACAAATGGTGTGTGAAACTCTGGATTCCTGGATTCGTGTGACATGTGACACTGAA	540
Db	481			
Qy	481	TTTTGGAA	TCCACAAATGGTGTGTGAAACTCTGGATTCCTGGATTCGTGTGACATGTGACACTGAA	540
Db	481			
Qy	541	ATAAGATTAC	GTGTTGCAATTATCTGTAATAACCTCCCTGCTCAATACCAAGAAAAAATTA	600
Db	541			
Qy	541	ATAAGATTAC	GTGTTGCAATTATCTGTAATAACCTCCCTGCTCAATACCAAGAAAAAATTA	600
Db	541			
Qy	601	AAAAAATCT	AGATTTAGGGTGAAGCTGACACTAGAAAGGCTCGAGGAAGATCAGCATGAT	660
Db	601			
Qy	601	AAAAAATCT	AGATTTAGGGTGAAGCTGACACTAGAAAGGCTCGAGGAAGATCAGCATGAT	660
Db	601			
Qy	661	AGGGTATCT	CCCCACTGTACTCCACAAATGTCCTCAATAGCTTGGAGATATCCTTAATAAGC	720
Db	661			
Qy	661	AGGGTATCT	CCCCACTGTACTCCACAAATGTCCTCAATAGCTTGGAGATATCCTTAATAAGC	720
Db	661			
Qy	721	GACAAATGAG	TTCAGTTCGACGCTTACACGCCGAGTGTGGTTATGGCTTTCGACGCTGAT	780
Db	721			
Qy	721	GACAAATGAG	TTCAGTTCGACGCTTACACGCCGAGTGTGGTTATGGCTTTCGACGCTGAT	780
Db	721			
Qy	781	CGTTGGACAG	AGTAGTACAGATACAGACGATGGAACCACTCGAACTAATCTTTTGAT	840
Db	781			
Qy	781	CGTTGGACAG	AGTAGTACAGATACAGACGATGGAACCACTCGAACTAATCTTTTGAT	840
Db	781			
Qy	841	TTTTTTCGAA	GAGATCTCAGTGAACAAGTTCAGGGTGAATGCCCTTCTTGACATGTG	900
Db	841			
Qy	841	TTTTTTCGAA	GAGATCTCAGTGAACAAGTTCAGGGTGAATGCCCTTCTTGACATGTG	900
Db	841			
Qy	901	GGTACAGCT	CTCTCTTATCATCCACATTCCTGAGAGTGGAAAGAGTGTCTGGAAATCTTT	960
Db	901			
Qy	901	GGTACAGCT	CTCTCTTATCATCCCATTCCTGAGAGTGGAAAGAGTGTCTGGAAATCTTT	960
Db	901			
Qy	961	ACTCTTCCCA	TCTCATGAGCAGAAATTTCCGGAAAAACAATAGGCCAAAGTGAGAGTTGACTAT	1020
Db	961			
Qy	961	ACTCTTCCCA	TCTCATGAGCAGAAATTTCCGGAAAAACAATAGGCCAAAGTGAGAGTTGACTAT	1020
Db	961			
Qy	1021	ATAAATTAAG	CCONTATACAGGATACAGTTGTGACATGAAATCTTCATTTTCCCAAGTAT	1080
Db	1021			
Qy	1021	ATAAATTAAG	CCONTATACAGGATACAGTTGTGACATGAAATCTTCATTTTCCCAAGTAT	1080
Db	1021			
Qy	1081	TGGAAAGCA	AGAAATACCAATTCGATGTGGCCCATCGAGGTGAGGAAACTCTCAACAACCT	1140
Db	1081			
Qy	1081	TGGAAAGCA	AGAAATACCAATTCGATGTGGCCCATCGAGGTGAGGAAACTCTCAACAACCT	1140
Db	1081			
Qy	1141	GCCCAGCT	GGCTTAAAGTTCAAGAAATATCTATGTGCTTCTTTAAGAAATGTCTGTAGTCAT	1200
Db	1141			
Qy	1141	GCCCAGCT	GGCTTAAAGTTCAAGAAATATCTATGTGCTTCTTTAAGAAATGTCTGTAGTCAT	1200
Db	1141			
Qy	1201	GGTGACGCT	TTGTAGAAATTTGACGTACACCTTTCAAGAGACTTTGTGCGCGTGGTATAT	1260
Db	1201			
Qy	1201	GGTGACGCT	TTGTAGAAATTTGACGTACACCTTTCAAGAGACTTTGTGCGCGTGGTATAT	1260
Db	1201			
Qy	1261	CATGATCTT	ACCTGTGTTTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTA	1320
Db	1261			
Qy	1261	CATGATCTT	ACCTGTGTTTGACTATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTA	1320
Db	1261			
Qy	1321	TTTTGAAAT	TCCAGTAAAGAAATTAACATTTGACCAACTCCAGTTGTTGAAAGCTCACTCAT	1380

[illegible]

Db 1039 TCATGAGCAGAAATTCGCGAAACAATAGCCAAAGTGAGGTGACTATATAATTATTA 1098
QY 1031 AGCCATTACCAAGGATACAGTTGTGACATGAATCTTCATTTTCAAGTATTTGAAGCCAA 1090
Db 1099 AGCCATTACCAAGGATACAGTTGTGACATGAATCTTCATTTTCAAGTATTTGAAGCCAA 1158
QY 1091 GAATACCAATTGGATTGGCCATCGAGGTGAGGAACTCTACAACTGCCCAGCTGG 1150
Db 1159 GAATACCAATTGGATTGGCCATCGAGGTGAGGAACTCTACAACTGCCCAGCTGG 1218
QY 1151 CTAAAGTTCAAGAAATACTATTCCTTTTAAAGAAATGCTGTAGTCATGTTGAGCCT 1210
Db 1219 CTAAAGTTCAAGAAATACTATTCCTTTTAAAGAAATGCTGTAGTCATGTTGAGCCT 1278
QY 1211 TTGTAGAAATTTGACGTACACCTTTTCAAGAGCTTTTGTGCCGTGGTATATCATGATCTTA 1270
Db 1279 TTGTAGAAATTTGACGTACACCTTTTCAAGAGCTTTTGTGCCGTGGTATATCATGATCTTA 1338
QY 1271 CCTGTGTTTGA CTATAGAAAAGAAATTTGATGCTGATGCCAGTTGAATTTTGAATTC 1330
Db 1339 CCTGTGTTTGA CTATAGAAAAGAAATTTGATGCTGATGCCAGTTGAATTTTGAATTC 1398
QY 1331 CAGTAAAGAAATTAACATTTGACCACTCCAGTTGTTAAAGCTCACCTCATGCTGACTGCAC 1390
Db 1399 CAGTAAAGAAATTAACATTTGACCACTCCAGTTGTTAAAGCTCACCTCATGCTGACTGCAC 1458
QY 1391 TCGAAATCTAAGGATCGGAAAGAAATCTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAATC 1450
Db 1459 TCGAAATCTAAGGATCGGAAAGAAATCTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAATC 1518
QY 1451 AGCCATTTCTCTTCTTAAGATGTTTGTAGAGTCTTTGCCAGAGATGTAGGGTTTAAACA 1510
Db 1519 AGCCATTTCTCTTCTTAAGATGTTTGTAGAGTCTTTGCCAGAGATGTAGGGTTTAAACA 1578
QY 1511 TTGAAATAAAATGGAATCTGCCAGCAAGGATGGAATGTGGATGGAATCTATCAACAT 1570
Db 1579 TTGAAATAAAATGGAATCTGCCAGCAAGGATGGAATGTGGATGGAATCTATCAACAT 1638
QY 1571 ATTTTCACATGAATCTGTTTGGATATAATTTTAAAGCTGTTTGTAGAAATTCCTGGGA 1630
Db 1639 ATTTTCACATGAATCTGTTTGGATATAATTTTAAAGCTGTTTGTAGAAATTCCTGGGA 1698
QY 1631 AGAGGAGAAATAGTGTTCCTTCATTTGATGCAATTTGCAATGTTGGCAAGC 1690
Db 1699 AGAGGAGAAATAGTGTTCCTTCATTTGATGCAATTTGCAATGTTGGCAAGC 1758
QY 1691 AGAACAAATATCGGATACCTATTTTAACTCAAGGAAATCTGAGATTTATCCTGAACCTCA 1750
Db 1759 AGAACAAATATCGGATACCTATTTTAACTCAAGGAAATCTGAGATTTATCCTGAACCTCA 1818
QY 1751 TGGACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTGCAAGTTTGAATAATCTAC 1810
Db 1819 TGGACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTGCAAGTTTGAATAATCTAC 1878
QY 1811 TGGGGAATAATGATACATCTGAGACTTGTCTGAGAAACCCATCTCTATATTCGAAGGCAA 1870
Db 1879 TGGGGAATAATGATACATCTGAGACTTGTCTGAGAAACCCATCTCTATATTCGAAGGCAA 1938
QY 1871 AAGCTAAGGGA CTAGTCATATTCCTGCTGGGTGATGATACCAATGATTCCTGAACAGAA 1930
Db 1939 AAGCTAAGGGA CTAGTCATATTCCTGCTGGGTGATGATACCAATGATTCCTGAACAGAA 1998
QY 1931 GGAATTTGAAGAACTTGGAGTTAAATGCTCTAATTTATGATAGGATATATGATGATGC 1990
Db 1999 GGAATTTGAAGAACTTGGAGTTAAATGCTCTAATTTATGATAGGATATATGATGATGC 2058
QY 1991 CTGAACCAACCAATATATTTTCAAGTGGAGCAATTTGGAAGCCCTGGAAGCAAGTATGCCAG 2050
Db 2059 CTGAACCAACCAATATATTTTCAAGTGGAGCAATTTGGAAGCCCTGGAAGCAAGTATGCCAG 2118
QY 2051 AGCTTAAAGAGCTGTTGTGCCATCTGTAGCCGCTTTGTTCCCTCATCTTTGTGTGGGG 2110

Db 2119 AGCTTAAAGAGCTGTTGTGTGCCACTGTGTAAGCCGCTTTGTCCCTCATCTTTGTGTGGGG 2178
QY 2111 AGCTCTGATATCCATGTCATGCCAACCGCATTTGATAACGTCGAGAAATCTTTAGTTTTTAT 2170
Db 2179 AGCTCTGATATCCATGTCATGCCAACCGCATTTGATAACGTCGAGAAATCTTTAGTTTTTAT 2238
QY 2171 TGCACAGAGGTCAATTTTGGGGCGTGCACCGCTGTTCTGGGTATTCATTTTTCACACTG 2230
Db 2239 TGCACAGAGGTCAATTTTGGGGCGTGCACCGCTGTTCTGGGTATTCATTTTTCACACTG 2298
QY 2231 AGCAATGTTGATCTATGCCCTTTTGGGCTTCTCAGTTCAATGAAGCAATAATGAAGTATTT 2290
Db 2299 AGCAATGTTGATCTATGCCCTTTTGGGCTTCTCAGTTCAATGAAGCAATAATGAAGTATTT 2358
QY 2291 AACTCTTTTCACTACAGTCTTCTGCAAGTATGCTATTTAAATTAATTTGGCCAGGTATAATG 2350
Db 2359 AACTCTTTTCACTACAGTCTTCTGCAAGTATGCTATTTAAATTAATTTGGCCAGGTATAATG 2418
QY 2351 CCAGTCAGTCTCTCTTATAGTGAGAAATTTATTTGGTTAGTAATAATAATTTTAAACTA 2410
Db 2419 CCAGTCAGTCTCTCTTATAGTGAGAAATTTATTTGGTTAGTAATAATAATTTTAAACTA 2478
QY 2411 AATATATAAATCTATAATGTTTAAA CATATGTTTCAATTAAGCATAGCACTTTGAAATTA 2470
Db 2479 AATATATAAATCTATAATGTTTAAA CATATGTTTCAATTAAGCATAGCACTTTGAAATTA 2538
QY 2471 CTATATAAATAGCTCATATTTTACATTCACAGCTTTTTCATTTTGTATCAGGTCGAAATCTTT 2530
Db 2539 CTATATAAATAGCTCATATTTTACATTCACAGCTTTTTCATTTTGTATCAGGTCGAAATCTTT 2598
QY 2531 AGCACTTAAAGGAAATGCACTATGCTATAATTTATACCTGACCAATGAAGAAATTAAGTACCTC 2590
Db 2599 AGCACTTAAAGGAAATGCACTATGCTATAATTTATACCTGACCAATGAAGAAATTAAGTACCTC 2658
QY 2591 AATATGATGATTTGCACTGGTGAATTTCCAACTGCACAAAATCTTTTGGCCATCTTGATAT 2650
Db 2659 AATATGATGATTTGCACTGGTGAATTTCCAACTGCACAAAATCTTTTGGCCATCTTGATAT 2718
QY 2651 AGTATTTTTTACATGGTGTGACATGACACACACCATTTTCATTCAGTATGAACTTTG 2710
Db 2719 AGTATTTTTTACATGGTGTGACATGACACACACCATTTTCATTCAGTATGAACTTTG 2778
QY 2711 AGGCTGCTGCATTTTTTCCATTTAAACCAACAGCTGAAAGTGAACCTCGAAATCTGTT 2770
Db 2779 AGGCTGCTGCATTTTTTCCATTTAAACCAACAGCTGAAAGTGAACCTCGAAATCTGTT 2838
QY 2771 TCATAAATCTTTCAAAAGTTGTTTACATCAATGTTTAAATTTTCAAAATGCTCGAGGTA 2830
Db 2839 TCATAAATCTTTCAAAAGTTGTTTACATCAATGTTTAAATTTTCAAAATGCTCGAGGTA 2898
QY 2831 ATTTTATGATATAATAATTTAGTAAGAAAGTATGTTTGCATACCTTAGTAGAATAGATC 2890
Db 2899 ATTTTATGATATAATAATTTAGTAAGAAAGTATGTTTGCATACCTTAGTAGAATAGATC 2958
QY 2891 ACAACATACAAATTTCAATTCAGTGCATGCTTTTAAAGTGTAAAGCATGAGATTTGATATGTT 2950
Db 2959 ACAACATACAAATTTCAATTCAGTGCATGCTTTTAAAGTGTAAAGCATGAGATTTGATATGTT 3018
QY 2951 TACTGTTAGGTCTTTCATCTGTTGGTCTAGGTGATGAGAAAGATGTCACAGACTGGA 3010
Db 3019 TACTGTTAGGTCTTTCATCTGTTGGTCTAGGTGATGAGAAAGATGTCACAGACTGGA 3078
QY 3011 CGTATTTTTTGGCTTAAAGAAAGGCTGTTTGTAGCGCTTTTAAATATGCTTTATTTT 3070
Db 3079 CGTATTTTTTGGCTTAAAGAAAGGCTGTTTGTAGCGCTTTTAAATATGCTTTATTTT 3138
QY 3071 TGTGCTCTCACTACCTATTACACATGTTTGTGCTTTTGGGGTTTGTTTGTATGTCGCTGT 3130
Db 3139 TGTGCTCTCTCACTACCTATTACACATGTTTGTGCTTTTGGGGTTTGTTTGTATGTCGCTGT 3198
QY 3131 GTTATACAGTATGTTTAAATTTCCATGCAAGAAATAAATATGTCGAAATCTTCAAAAAAAA 3190
Db 3199 GTTATACAGTATGTTTAAATTTCCATGCAAGAAATAAATATGTCGAAATTTTCAAAAAAAA 3258

10/647855

1527 AGCCATTTCTTCTTAAGATGGTTTTAGAGTCTTTGCCAGAGAAGTGTAGGGTTAAACA 1586
1511 TTGAATAAATAGATCTGCAGCAAGGAGTGAATGTGGATGGTAACTTATCAACAT 1570
1587 TTGAATAAATAGATCTGCAGCAAGGAGTGAATGTGGATGGTAACTTATCAACAT 1646
1571 ATTTTGACATGAATCTGTTTTGGATATAAATTTTAAATACTGTTTTAGAAAATCTGGGA 1630
1647 ATTTTGAATGAATCTGTTTTGGATATAAATTTTAAATACTGTTTTAGAAAATCTGGGA 1706
1631 AGAGGAGAAATAGTGTCTTTCTTCAATTTGATGAGATATTTTGCACAATGGTTGCGCAAAAGC 1690
1707 AGAGGAGAAATAGTGTCTTTCTTCAATTTGATGAGATATTTTGCACAATGGTTGCGCAAAAGC 1766
1691 AGAACAAATATCCGATATATTTTAACTCAAGGAAATCTGAGATTAATCTGCAACTCA 1750
1767 AGAACAAATATCCGATATATTTTAACTCAAGGAAATCTGAGATTAATCTGCAACTCA 1826
1751 TGGACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTGCAAGTTTGAATAATCTAC 1810
1827 TGGACCTCAGATCTCGGACAAACCCCATTTGCAATGAGCTTTGCAAGTTTGAATAATCTAC 1886
1811 TGGGATAAATGTACATCTGAAGACTTTGCTCAGAAAACCCATCTCTATATTTCAAGAGGCAA 1870
1887 TGGGATAAATGTACATCTGAAGACTTTGCTCAGAAAACCCATCTCTATATTTCAAGAGGCAA 1946
1871 AAGCTAAGGAGTGTATCTATCTCTGGGGTGTATGATACCAATGATCTCTGAAAACAGAA 1930
1947 AAGCTAAGGAGTGTATCTATCTCTGGGGTGTATGATACCAATGATCTCTGAAAACAGAA 2006
1931 GGAATTTGAAGAACTTGGAGTTAATGCTTAATTTATGATAGGATATATGATTTGATGTC 1990
2007 GGAATTTGAAGAACTTGGAGTTAATGCTTAATTTATGATAGGATATATGATTTGATGTC 2066
1991 CTGAACAAACCAATATATTTTCAAGTGGAGCAATTTGGAACGGCTGGAAGCAGGAATTTGCCAG 2050
2067 CTGAACAAACCAATATATTTTCAAGTGGAGCAATTTGGAACGGCTGGAAGCAGGAATTTGCCAG 2126
2051 AGCTTAAAGAGCTGTTGTGTCCCATCTGTTAGCGCTTTTGTTCCTCATCTTTGTGTGGGG 2110
2127 AGCTTAAAGAGCTGTTGTGTCCCATCTGTTAGCGCTTTTGTTCCTCATCTTTGTGTGGGG 2186
2111 AGCTGATATCAATGTGATGTCACCGCAATTTGATTAAGTGGAGATCTGATTTTAT 2170
2187 AGCTGATATCAATGTGATGTCACCGCAATTTGATTAAGTGGAGATCTGATTTTAT 2246
2171 TGCACAGAGTCAATTTTGGGGCGTGCACCGCTGTTCTGGGTATTCATTTTTCATCACTG 2230
2247 TGCACAGAGTCAATTTTGGGGCGTGCACCGCTGTTCTGGGTATTCATTTTTCATCACTG 2306
2231 AGCATTTGATCTATGCTTTTGGGCTTCTCAGTTCAATGAAGCAATTAAGTATTT 2290
2307 AGCATTTGATCTATGCTTTTGGGCTTCTCAGTTCAATGAAGCAATTAAGTATTT 2366
2291 AACTCTTTCACTACAGTCTTGAAGTATGCTATTTAAATTAATTTACTTGGCCAGGTAAATG 2350
2367 AACTCTTTCACTACAGTCTTGAAGTATGCTATTTAAATTAATTTACTTGGCCAGGTAAATG 2426
2351 CCAGTCAGTCTCTTTATAGTGAATAATTTATTTGTTAGTATAATAATTTTAAACTA 2410
2427 CCAGTCAGTCTCTTTATAGTGAATAATTTATTTGTTAGTATAATAATTTTAAACTA 2486
2411 AATATATAAATCTATATGTTAAACAATGTTTCAATTAAGCATAGCACTTTGAAATTA 2470
2487 AATATATAAATCTATATGTTAAACAATGTTTCAATTAAGCATAGCACTTTGAAATTA 2546
2471 CTATATAAATAGTCTATAATTTCACTACAGCTTTTCAATTTGATCAGGCTCTGAATCTTT 2530
2547 CTATATAAATAGTCTATAATTTCACTACAGCTTTTCAATTTGATCAGGCTCTGAATCTTT 2606
2531 AGCACTTAAAGAAATGCTATGCAATTAATTAACCTGACCATGAAAATAATAGTACCTC 2590
2607 AGCACTTAAAGAAATGCTATGCAATTAATTAACCTGACCATGAAAATAATAGTACCTC 2666

2591 AAATGATCATGATTTGACATGGTGAATTTCCAACTGCAAAAATCTTTGTGCCATCTTGTATAT 2650
2667 AAATGATCATGATTTGACATGGTGAATTTCCAACTGCAAAAATCTTTGTGCCATCTTGTATAT 2726
2651 AGGTATTTTTTATCATGGGTGATGCAACAACCAATTTTCAATTCAGTATGAACCTTG 2710
2727 AGGTATTTTTTATCATGGGTGATGCAACAACCAATTTTCAATTCAGTATGAACCTTG 2786
2711 AGGCTGCTGCCATTTTCCATTTTAAACCAACCAAGCTGAACCTCGAAAATCTTGT 2770
2787 AGGCTGCTGCCATTTTCCATTTTAAACCAACCAAGCTGAACCTCGAAAATCTTGT 2846
2771 TCATAAATCTTTCAAAAGTTGTTTTTACATCAATGTTTAAATTTCAAAATGCTCGAGGTA 2830
2847 TCATAAATCTTTCAAAAGTTGTTTTTACATCAATGTTTAAATTTCAAAATGCTCGAGGTA 2906
2831 ATTTAATGATATAAATATTTAGTAAGAAAAGTATGATTTGATATCTTAGTAGAATAGATC 2890
2907 ATTTAATGATATAAATATTTAGTAAGAAAAGTATGATTTGATATCTTAGTAGAATAGATC 2966
2891 ACAACATACAAATTTCAATTCAGTGCATGCTTTTAGGTGTTAAGCATGAGATTTGATATGTT 2950
2967 ACAACATACAAATTTCAATTCAGTGCATGCTTTTAGGTGTTAAGCATGAGATTTGATATGTT 3026
2951 TACTGTTAGTCTCTGATCTGCTGCTGATGATGATGAGTAAGTAAGTGTCAAGGACTGGA 3010
3027 TACTGTTAGTCTCTGATCTGCTGCTGATGATGATGAGTAAGTAAGTGTCAAGGACTGGA 3086
3011 CGTATTTTGTGCTCAAAAAAAGCGTGTGTTAGGCGTTTTTAAATATGCTTATTTTG 3070
3087 CGTATTTTGTGCTCAAAAAAAGCGTGTGTTAGGCGTTTTTAAATATGCTTATTTTG 3146
3071 TGTGCTCTCACTACCTATATACACATGTTGCTTCTGCGGTTGTTTGTATGCTGCTGT 3130
3147 TGTGCTCTCACTACCTATATACACATGTTGCTTCTGCGGTTGTTTGTATGCTGCTGT 3206
3131 GTTATACAGTAGTTAAATTTCCATGCAAGAAAATAAATGCTCTGAATTTCTCAA 3184
3207 GTTATACAGTAGTTAAATTTCCATGCAAGAAAATAAATGCTCTGAATTTCTCAT 3260

RESULT 5
US-10-357-930-23730
; Sequence 23730, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23730


```

: LENGTH: 3499
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: misc feature
: LOCATION: 3497..3491, 3492, 3493, 3494, 3495, 3496, 3497, 3498, 3499
: OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23730

```

Query Match 98.7%; Score 3164.4; DB 20; Length 3499;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3168; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	11	CGCTCGGCGGAGGCACGGACGGCGGGCGCCGGTACTCTGCTCCCGGGTCTCGCTCTCTC	146
Db	87	GAGCGTGGCGAGGCACGGACGGCGGGCGCCGGTACCTCTGCGCGGGTCTCGCTCTCTC	146
Qy	71	GGGCGGGCGGCGGACGCGGACCTGCGGACTAGCGAAACCCGAGACGACATCATAA	130
Db	147	GGGCGGGCGGCGGACGCGGACCTGCGGACTAGCGAAACCCGAGACGACATCATAA	206
Qy	131	ATAAATCCATCAGAATGACACCTTCTCAGGTGTGCTTTGAAATAAGAGGAACCTCTTTTAC	190
Db	207	ATAAATCCATCAGAATGACACCTTCTCAGGTGTGCTTTGAAATAAGAGGAACCTCTTTTAC	266
Qy	191	CAGGAGAACTTTTTCGATATGTGGAAGCTGTGATGCTTTTGGGAACTCGGAATCCTCAAA	250
Db	267	CAGGAGAACTTTTTCGATATGTGGAAGCTGTGATGCTTTTGGGAACTCGGAATCCTCAAA	326
Qy	251	ATGCTGTGGCTCTTCTCCAGAGATGACACAGCTGGAAGACGTCATGGAAGCAACCA	310
Db	327	ATGCTGTGGCTCTTCTCCAGAGATGACACAGCTGGAAGACGTCATGGAAGCAACCA	386
Qy	311	TTGTACTCAGTAGAGGAGTATCAGTTTCAGTATCGCTACTTTCAAAGGGTACTTTTTAGAAC	370
Db	387	TTGTACTCAGTAGAGGAGTATCAGTTTCAGTATCGCTACTTTCAAAGGGTACTTTTTAGAAC	446
Qy	371	CAAGAATATCGGTGGTCCATGTCAAGTGATAGTTTCAAAAGTGGAGACTCATCTTCAAC	430
Db	447	CAAGACTATCGGTGGTCCATGTCAAGTGATAGTTTCAAAAGTGGAGACTCATCTTCAAC	506
Qy	431	CAGGATCAATAACCCCTTTAGAAAGCGAAATATATTGACGATGACCAATTTGGGAATCC	490
Db	507	CAGGATCAATAACCCCTTTAGAAAGCGAAATATATTGACGATGACCAATTTGGGAATCC	566
Qy	491	ACAATGGTGTTCGAAACTCTGGATTCTGATGCGCTGACATGTGCAAGCTGGAATAAGATTAC	550
Db	567	ACAATGGTGTGAAACTCTGGATTCTGATGCGCTGACATGTGCAAGCTGGAATAAGATTAC	626
Qy	551	GTTTGCAATTATTCTGAAAAACCTCTGTGTCAATAACCAAGAAAAATTTAAAAAATCTTA	610
Db	627	GTTTGCAATTATTCTGAAAAACCTCTGTGTCAATAACCAAGAAAAATTTAAAAAATCTTA	686
Qy	611	GATTTAGGTGAACTGACATAGAAAGCCCTGGAGGAAGATGACGATGATAGGGTATCTTC	670
Db	687	GATTTAGGTGAACTGACATAGAAAGCCCTGGAGGAAGATGACGATGATAGGGTATCTTC	746
Qy	671	CCACTGTACTCCACAAAATGTCCAAATAGCTTGGAGATATCCTTAAATAAGCGCAATGAGT	730
Db	747	CCACTGTACTCCACAAAATGTCCAAATAGCTTGGAGATATCCTTAAATAAGCGCAATGAGT	806
Qy	731	TCAAGTGCAGGCAATTCACAGCGGAGTGTGGTATGTGCTTGACGCTGATGCTTTGGACAG	790
Db	807	TCAAGTGCAGGCAATTCACAGCGGAGTGTGGTATGTGCTTGACGCTGATGCTTTGGACAG	866
Qy	791	AGTAGAGCATACAGACGATGGAAACCGATAACTTGGAACTAATCTTTGATTTTTTTCGAAG	850
Db	867	AGTAGAGCATACAGACGATGGAAACCGATAACTTGGAACTAATCTTTGATTTTTTTCGAAG	926
Qy	851	AAGATCTCAGTGAGCAGCTAGTTTCAGGGTGTATGTCCTTCTGACATGTGGGTACAGCTT	910
Db	927	AAGATCTCAGTGAGCAGCTAGTTTCAGGGTGTATGTCCTTCTGACATGTGGGTACAGCTT	986

Qy	911	GTCTCTTATCATCCACCATTGCTGAGAGTGGAAAGAGTGTCTGGAATTCCTTACTCTTCCCCA	970
Db	987	GTCTCTTATCATCCACCATTGCTGAGAGTGGAAAGAGTGTCTGGAATTCCTTACTCTTCCCCA	1046
Qy	971	TCATGACGAAATTCCTCCGAAAAACAATATGGCAAAAGTGAGAGTTGACTATATAATTATTATTA	1030
Db	1047	TCATGACGAGAAATTCCTCCGAAAAACAATATGGCAAAAGTGAGAGTTGACTATATAATTATTATT	1106
Qy	1031	AGCCATTACAGGATACAGTTGTGACACATGAAATCTTCAATTTTCCAAAGTATGTGAAGCCAA	1090
Db	1107	AGCCATTACAGGATACAGTTGTGACATGAAATCTTCAATTTTCCAAAGTATGTGAAGCCAA	1166
Qy	1091	GAATACCATTTGGATGTGGCCATCGAGGTCCAGGAACCTACAAACACTGCCCGAGCTGG	1150
Db	1167	GAATACCATTTGGATGTGGCCATCGAGGTTCAGGAACTCTCAACAACTGCCCGAGCTGG	1226
Qy	1151	CTAAAGTTCAAGAAATACATTAATGCTTCTTTTAAAGAAATGCTGCTAGTCACTGGTGCAGCCT	1210
Db	1227	CTAAAGTTCAAGAAATACATTAATGCTTCTTTTAAAGAAATGCTGCTAGTCACTGGTGCAGCCT	1286
Qy	1211	TTGTAGAAATTTGACGTAACCTTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTA	1270
Db	1287	TTGTAGAAATTTGACGTACACCTTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTA	1346
Qy	1271	CTGTGTTGTTGACTATGAAAAAGAAATTTGATGCTGATTCAGTTGAAATTAATTGAAATTC	1330
Db	1347	CTGTGTTGTTGACTATGAAAAAGAAATTTGATGCTGATTCAGTTGAAATTAATTGAAATTC	1406
Qy	1331	CAGTAAAGAAATTAACATTTGACCACTCCAGTTGTTTAAAGCTCACATCTGCTGACATGCGAC	1390
Db	1407	CAGTAAAGAAATTTAAACATTTGACCACTCCAGTTGTTTAAAGCTCACATCTGCTGACATGCGAC	1466
Qy	1391	TGAAATCTAAGGATCGGAAAGAAATCTGTGTTTCCAGGAGGAAAAATTCCTTTTCAGAAAAATC	1450
Db	1467	TGAAATCTAAGGATCGGAAAGAAATCTGTGTTTCCAGGAGGAAAAATTCCTTTTCAGAAAAATC	1526
Qy	1451	AGCCATTCCTTCTCTTAAGATGGTTTTAGAGTCTTTGCCAGAAAGATGTAGGGTTTAAACA	1510
Db	1527	AGCCATTCCTTCTCTTAAGATGGTTTTAGAGTCTTTGCCAGAAAGATGTAGGGTTTAAACA	1586
Qy	1511	TTGAAATAAAATCGATCTGCCAGCAAGGATGGAATGTGGAGTGTAACTTATCAACAT	1570
Db	1587	TTGAAATAAAATCGATCTGCCAGCAAGGATGGAATGTGGAGTGTAACTTATCAACAT	1646
Qy	1571	ATTTTGACATGAATCTGTTTTTGGATATAATTTTAAAAAATCTGTTTTAGAAAAATTCCTGGGA	1630
Db	1647	ATTTTGACATGAATCTGTTTTTGGATATAATTTTAAAAAATCTGTTTTAGAAAAATTCCTGGGA	1706
Qy	1631	AGAGGAGATAGTGTGTTTTCTTCATTTTGATGACAGATATTTTGCACAATGGTTCGGCAAAAGC	1690
Db	1707	AGAGGAGATAGTGTGTTTTCTTCATTTTGATGACAGATATTTTGCACAATGGTTCGGCAAAAGC	1766
Qy	1691	AGAACAAATATCCGATACATATTTTAACTTCAAGGAAAACTGAGATTATCTCTGAACTCA	1750
Db	1767	AGAACAAATATCCGATACATATTTTAACTTCAAGGAAAACTGAGATTATCTCTGAACTCA	1826
Qy	1751	TGACCTCTCAGATCTCCGACAAACCCCATTTGCAATGAGCTTTGCACAGTTTGGAAAAATCTAC	1810
Db	1827	TGACCTCTCAGATCTCCGACAAACCCCATTTGCAATGAGCTTTGCACAGTTTGGAAAAATCTAC	1886
Qy	1811	TGGGGATAAATGTACATACCTGAAAGACTTGTCTCAGAAACCCCATCTTATTTCAAGAGGCAA	1870
Db	1887	TGGGGATAAATGTACATACCTGAAAGACTTGTCTCAGAAACCCCATCTTATTTCAAGAGGCAA	1946
Qy	1871	AACTAAGGACTAGTGCATATTTCTGCTGGGGTGAATGATACCAATGATCTCTGAAAAACAGAA	1930
Db	1947	AACTAAGGACTAGTGCATATTTCTGCTGGGGTGAATGATACCAATGATCTCTCTGAAAAACAGAA	2006
Qy	1931	GGAAATTTGAAGGAATCTGGAGTTAAATGGTCTAAATTTTATGATAGGATATATGATTTGATGC	1990
Db	2007	GGAAATTTGAAGGAATCTGGAGTTAAATGGTCTAAATTTTATGATAGGATATATGATTTGATGC	2066
Qy	1991	CTGAACCAACAAATATATTCCAAGTGGAGCAATTTGGAAACGCTCTGAAGACAGGAATTCGCCAG	2050

Db	2067	CTGAACCAACCAATATATTCCTCAAGTGGAGCAATTTGGAACCGCTGAAGCAGGAATTCGCAG	2126	
Qy	2051	AGCTTAAGAGCTGTTTGTGTCCCACTGTTAGCCGCTTTGTTCCTCATCTTTGTGTGGGG	2110	
Db	2127	AGCTTAAGAGCTGTTTGTGTCCCACTGTTAGCCGCTTTGTTCCTCATCTTTGTGTGGGG	2186	
Qy	2111	AGCTGATATCCATGTGTGATGCCAACGGCATTGTAATAGTGGAGATGCTTAGTTTAT	2170	
Db	2187	AGTCTGATATCCATGTGTGATGCCAACGGCATTGTAATAGTGGAGATGCTTAGTTTAT	2246	
Qy	2171	TGCACAGAGCTCATTTTGGGGGGCTGCACCGCTCTCTGGGTATTCATTTTTCATCACTG	2230	
Db	2247	TGCACAGAGCTCATTTTGGGGGGCTGCACCGCTCTCTGGGTATTCATTTTTCATCACTG	2306	
Qy	2231	AGCATTTGTTGATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAATAATGAAGTATTT	2290	
Db	2307	AGCATTTGTTGATCTATGCTTTTGGGCTTCTCAGTTCATTAAGCAATAATGAAGTATTT	2366	
Qy	2291	AACTCTTTCACTACAGTCTTTCGAAGTAGCTATTTAAATTTACTTGGCCAGGTATTAATG	2350	
Db	2367	AACTCTTTCACTACAGTCTTTCGAAGTAGCTATTTAAATTTACTTGGCCAGGTATTAATG	2426	
Qy	2351	CCAGTCAGTCTCTTATAGTGAGAAATTTATTTGGTTAGTAAATAATATTTTAAACTA	2410	
Db	2427	CCAGTCAGTCTCTTATAGTGAGAAATTTATTTGGTTAGTAAATAATATTTTAAACTA	2486	
Qy	2411	ATATATAAATCTATAATGTTTAAACATATGTTTCAATTAAGCAATAGACCTTTGAAATTA	2470	
Db	2487	ATATATAAATCTATAATGTTTAAACATATGTTTCAATTAAGCAATAGACCTTTGAAATTA	2546	
Qy	2471	CTATATAAATAGTCTCATATTTACATTAAGTCTTTCATTTGATCAGTCTGAAATCTTT	2530	
Db	2547	CTATATAAATAGTCTCATATTTACATTAAGTCTTTCATTTGATCAGTCTGAAATCTTT	2606	
Qy	2531	AGCACTTAAGGAAATGATGATGATTAATTAACCTGACCATGAAATAAATAGTACCTC	2590	
Db	2607	AGCACTTAAGGAAATGATGATGATTAATTAACCTGACCATGAAATAAATAGTACCTC	2666	
Qy	2591	AAATGATCATTTGACATGCTGATTTTCAACTGCAAAATCTTTGTCATCTTGTATAT	2650	
Db	2667	AAATGATCATTTGACATGCTGATTTTCAACTGCAAAATCTTTGTCATCTTGTATAT	2726	
Qy	2651	AGGTATTTTACATGGGTGACATGCACACACCACTTTTCATTCAGTAGTAACCTTG	2710	
Db	2727	AGGTATTTTACATGGGTGACATGCACACACCACTTTTCATTCAGTAGTAACCTTG	2785	
Qy	2711	AGGCTGCTGCATTTTCCACTTAACCAACCGAGCTGAAGTGAACCTCGAACTGTT	2770	
Db	2787	AGGCTGCTGCATTTTCCACTTAACCAACCGAGCTGAAGTGAACCTCGAACTGTT	2846	
Qy	2771	TCATAAATCTTCAAAAGTTGTTTATACATCAATGTTTAAATTTTCAAAATGCTGCAGGTA	2830	
Db	2847	TCATAAATCTTCAAAAGTTGTTTATACATCAATGTTTAAATTTTCAAAATGCTGCAGGTA	2906	
Qy	2831	ATTTAATGATAAATAATAGTAAGAAAGTATGTAATGCACTTAGTAGAATAGATC	2890	
Db	2907	ATTTAATGATAAATAATAGTAAGAAAGTATGTAATGCACTTAGTAGAATAGATC	2966	
Qy	2891	ACAACATACAAATTCATTCAGTGCATGCTTTAGGTGTTAAGCATGAGATTTGATATGTT	2950	
Db	2967	ACAACATACAAATTCATTCAGTGCATGCTTTAGGTGTTAAGCATGAGATTTGATATGTT	3026	
Qy	2951	TACTGTTAGTCTTGCATCTGTTGCTAGTGAATGATGAGAAAGATGCAAGGACTGGA	3010	
Db	3027	TACTGTTAGTCTTGCATCTGTTGCTAGTGAATGATGAGAAAGATGCAAGGACTGGA	3086	
Qy	3011	CGTATTTTGTGCTCAAAAAAAGGCTGTTTGTAGGCGTTTAAATATGCTTATTTTG	3070	
Db	3087	CGTATTTTGTGCTCAAAAAAAGGCTGTTTGTAGGCGTTTAAATATGCTTATTTTG	3146	
Qy	3071	TGCTCTCTCACTACCTATTACACATGTTGCTTTTGGGTTGTTTGTATGTCGTGT	3130	
Db	3147	TGTGTCTCTCACTACCTATTACACATGTTGCTTTTGGGTTGTTTGTATGTCGTGT	3206	
Qy	3131	GTTATACAGTAGTTTAAATTTCCATGCAGAAAAATAAATGCTCTGAATTCCTCAA	3184	
Db	3207	GTTATACAGTAGTTTAAATTTCCATGCAGAAAAATAAATGCTCTGAATTTTCTATA	3260	
RESULT 6				
US-10-357-930-29632				
; Sequence 29632, Application US/10357930				
; Publication No. US20040259086A1				
; GENERAL INFORMATION:				
; APPLICANT: Schlegel, Robert				
; APPLICANT: Endege, Wilson				
; APPLICANT: Monahan, John				
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR				
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF				
; TITLE OF INVENTION: HUMAN PROSTATE CANCER				
; FILE REFERENCE: MRI-007BCN				
; CURRENT APPLICATION NUMBER: US/10/357,930				
; PRIOR FILING DATE: 2003-02-04				
; PRIOR APPLICATION NUMBER: 09/785,276				
; PRIOR FILING DATE: 2003-02-16				
; PRIOR APPLICATION NUMBER: 60/183,319				
; PRIOR FILING DATE: 2000-02-17				
; PRIOR APPLICATION NUMBER: 60/189,862				
; PRIOR FILING DATE: 2000-03-16				
; PRIOR APPLICATION NUMBER: 60/207,454				
; PRIOR FILING DATE: 2000-05-25				
; PRIOR APPLICATION NUMBER: 60/211,314				
; PRIOR FILING DATE: 2000-06-09				
; PRIOR APPLICATION NUMBER: 60/219,007				
; PRIOR FILING DATE: 2000-07-18				
; PRIOR APPLICATION NUMBER: 60/255,281				
; PRIOR FILING DATE: 2000-12-13				
; NUMBER OF SEQ ID NOS: 62232				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 29632				
; LENGTH: 3499				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: misc feature				
; LOCATION: 3490_3491, 3492, 3493, 3494, 3495, 3496, 3497, 3498, 3499				
; OTHER INFORMATION: n = A,T,C or G				
US-10-357-930-29632				
Query Match				
Best Local Similarity				
Matches 3168; Conservative				
Score 3164.4; DB 20; Length 3499;				
Pred. No. 0;				
Mismatches 6; Indels 0; Gaps 0;				
Qy	11	CGTCCGGGCGAGCAGGACGGGCGCGGCTACCTCTGCGCGGCTCCTCGCTCTC	70	
Db	87	GAGCGCTGGCGAGGACGAGCGGCGCGGCTACCTCTGCGCGGCTCCTCGCTCTC	146	
Qy	71	GGGCGGCGCGCGGACGCGGACCTCGGACCTAGCGAAACCGGAGCAGCATCATATAA	130	
Db	147	GGGCGGCGCGCGGAGGACGCGGACCTCGGACCTAGCGAAACCGGAGCAGCATCATATAA	206	
Qy	131	ATAAATCCATCAGAAATGACACCTTCTCAGGTTGCTTTGAAATTAAGAGGAACCTTTTAC	190	
Db	207	ATAAATCCATCAGAAATGACACCTTCTCAGGTTGCTTTGAAATTAAGAGGAACCTTTTAC	266	
Qy	191	CAGGAGAAATTTTTCGATATGTGGAAGCTGTGATGCTTTGGGAACTGGAACTCCTCAA	250	
Db	267	CAGGAGAAATTTTTCGATATGTGGAAGCTGTGATGCTTTGGGAACTGGAACTCCTCAA	326	
Qy	251	ATGCTGTGCTCTTCTTCCAGAGATGACACAGGTGAAGCATGCTATGGAAGCAACCA	310	
Db	327	ATGCTGTGCTCTTCTTCCAGAGATGACACAGGTGAAGCATGCTATGGAAGCAACCA	386	
Qy	311	TTGTAATCAGTAGAGGAGTATCAGTTAGTATCGCTCTTCAAAGGGTACTTTTTAGAAC	370	

Db 387 TTGTACTCAGTAGAGGATATCAGTTCAAGTTCGCTACTTCAAGGGTACTTTTATGAAC 446
Qy 371 CAAAGACTATCGGTGTCCATGTCAGTTCAGTTCAGAGTGGGAGACTCATCTCAAC 430
Db 447 CAAAGACTATCGGTGTCCATGTCAGTTCAGTTCAGAGTGGGAGACTCATCTCAAC 506
Qy 431 CACGATCAATACCCCTTTAGAAAGCGAAATTAATTTAGCAGTGCAGCAATTTGGATCC 490
Db 507 CACGATCAATACCCCTTTAGAAAGCGAAATTAATTTAGCAGTGCAGCAATTTGGATCC 566
Qy 491 ACAATGGTGTGAACCTCTGGATTCTGGATGCTGCATGTCAGACTGCAATTAAGATTAC 550
Db 567 ACAATGGTGTGAACCTCTGGATTCTGGATGCTGCATGTCAGACTGCAATTAAGATTAC 626
Qy 551 GTTTGCAATTAATCTGAAAAACCTCTGTGTCAATACCAAGAAAAAATTAAGAAATCTA 610
Db 627 GTTTGCAATTAATCTGAAAAACCTCTGTGTCAATACCAAGAAAAAATTAAGAAATCTA 686
Qy 611 GATTTAGGGTGAAGCTGACACTAGAGGCTTGGAGGAAGATGACGATGATAGGATCTC 670
Db 687 GATTTAGGGTGAAGCTGACACTAGAGGCTTGGAGGAAGATGACGATGATAGGATCTC 746
Qy 671 CCACTGTACTCCACAAAAATGTCGAATAGCTTTGGAGATATCTTAATAAGCGACAATGACT 730
Db 747 CCACTGTACTCCACAAAAATGTCGAATAGCTTTGGAGATATCTTAATAAGCGACAATGACT 806
Qy 731 TCAAGTGCAGGCAATTCACAGCGGAGTGTGGTTATGGCTTGCAGCTGATCGTTGGACAG 790
Db 807 TCAAGTGCAGGCAATTCACAGCGGAGTGTGGTTATGGCTTGCAGCTGATCGTTGGACAG 866
Qy 791 AGTACAGCATACAGAGTGGAAACAGATAACTTGAACCTTAATCTTTGATTTTTTCGAAG 850
Db 867 AGTACAGCATACAGAGTGGAAACAGATAACTTGAACCTTAATCTTTGATTTTTTCGAAG 926
Qy 851 AAGATCTCAGTCAGCAGCTAGTTTCAGGGTGTGSCCTTCTGACATGATGGGTGACAGTT 910
Db 927 AAGATCTCAGTCAGCAGCTAGTTTCAGGGTGTGSCCTTCTGACATGATGGGTGACAGTT 986
Qy 911 GTCTCTTATCTCACCAATTCGTGAGTGGAAAGAGTGTGGAATTCCTTACTCTTCCCA 970
Db 987 GTCTCTTATCATCCACCAATTCGTGAGTGGAAAGAGTGTGGAATTCCTTACTCTTCCCA 1046
Qy 971 TCATGAGCAGAAATTCCTCGGAAAAAATAAGGCAAGTGAAGTGAATTAATTA 1030
Db 1047 TCATGAGCAGAAATTCCTCGGAAAAAATAAGGCAAGTGAAGTGAATTAATTA 1106
Qy 1031 AGCCATTACAGGATACAGTTGACATGAATCTTCAATTTCCAGTATTCGAAGCCAA 1090
Db 1107 AGCCATTACAGGATACAGTTGACATGAATCTTCAATTTCCAGTATTCGAAGCCAA 1166
Qy 1091 GAATACCAATGGATGTGGCCATCGAGGTGCAGGAAACTCTACAACTGCCAGCTGG 1150
Db 1167 GAATACCAATGGATGTGGCCATCGAGGTGCAGGAAACTCTACAACTGCCAGCTGG 1226
Qy 1151 CTAAGTTCAGAAATACTATTTGCTTTTAAGAAATGCTGTAGTCATGTCAGGCT 1210
Db 1227 CTAAGTTCAGAAATACTATTTGCTTTTAAGAAATGCTGTAGTCATGTCAGGCT 1286
Qy 1211 TTGTAGAATTTGACGTPACACCTTCAAGGACTTTGTGCGCGGTGATATCATGATCTTA 1270
Db 1287 TTGTAGAATTTGACGTPACACCTTCAAGGACTTTGTGCGCGGTGATATCATGATCTTA 1346
Qy 1271 CTTGTGTTTGAATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTAATTTGAAATTC 1330
Db 1347 CTTGTGTTTGAATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTAATTTGAAATTC 1406
Qy 1331 CAGTAAAAAGAAATTAATTTGAACCAACTCCAGTGTGTTAAAGCTCACTCATGTCGAC 1390
Db 1407 CAGTAAAAAGAAATTAATTTGAACCAACTCCAGTGTGTTAAAGCTCACTCATGTCGAC 1466
Qy 1391 TGAATCTTAAGATCGGAAGATCTGTGTTTCAGGAGGAAATTCCTTTTCAGAAATC 1450
Db 1467 TGAATCTTAAGATCGGAAGAAATCTGTGTTTCAGGAGGAAATTCCTTTTCAGAAATC 1526

Qy 1451 AGCCATTTCTTCTCTTAAGATGGTTTATAGAGTCTTTGCCAGAAAGTGTAGGTTTAAACA 1510
Db 1527 AGCCATTTCTTCTCTTAAGATGGTTTATAGAGTCTTTGCCAGAAAGTGTAGGTTTAAACA 1586
Qy 1511 TTGAATTAATTAAGTGTCCAGCAAGGGATGGATCTGGATGGTAACTTATCAACAT 1570
Db 1587 TTGAATTAATTAAGTGTCCAGCAAGGGATGGATCTGGATGGTAACTTATCAACAT 1646
Qy 1571 ATTTTGACATGAATCTGTTTTTGGATATAATTTTAAAAAATCTTTTAAAAAATCTGGGA 1630
Db 1647 ATTTTGACATGAATCTGTTTTTGGATATAATTTTAAAAAATCTTTTAAAAAATCTGGGA 1706
Qy 1631 AGAGGAGAAATAGTGTCTTCAATTTGATGCAAGATTAATTTGCAAAATGTTTCGCAAAAAGC 1690
Db 1707 AGAGGAGAAATAGTGTCTTCAATTTGATGCAAGATTAATTTGCAAAATGTTTCGCAAAAAGC 1766
Qy 1691 AGAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGAACCTCA 1750
Db 1767 AGAACAAATATCCGATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGAACCTCA 1826
Qy 1751 TGGACCTCAGATCTCGGACAAACCCCAATTTGCAATGAGCTTTGACAGTTTGAATCTTAC 1810
Db 1827 TGGACCTCAGATCTCGGACAAACCCCAATTTGCAATGAGCTTTGACAGTTTGAATCTTAC 1886
Qy 1811 TGGGATAAATGTACATACTGAAGACTTTGCTCAGAAACCCCAATCTTATATTTCAAGAGGCAA 1870
Db 1887 TGGGATAAATGTACATACTGAAGACTTTGCTCAGAAACCCCAATCTTATATTTCAAGAGGCAA 1946
Qy 1871 AAGCTTAAGGACTAGTCTATATTTGCTGGGTGATGATACCAATGATCTTGAANAACAGAA 1930
Db 1947 AAGCTTAAGGACTAGTCTATATTTGCTGGGTGATGATACCAATGATCTTGAANAACAGAA 2006
Qy 1931 GGAATTCAGAGAACTTCGGAGTTAATGCTTAAATTTATGATAGGATATATGATTCGATGC 1990
Db 2007 GGAATTCAGAGAACTTCGGAGTTAATGCTTAAATTTATGATAGGATATATGATTCGATGC 2066
Qy 1991 CTGAACAAACCAAAATATTTCAAGTGGAGCAATTTGGAACGCTTGAAGCAGGAATTTGCCAG 2050
Db 2067 CTGAACAAACCAAAATATTTCAAGTGGAGCAATTTGGAACGCTTGAAGCAGGAATTTGCCAG 2126
Qy 2051 AGCTTAAGAGCTGTTGTCCTGCTAGTTCAGCGCTTTGTTCCCTCATCTTTTGTGTGGG 2110
Db 2127 AGCTTAAGAGCTGTTGTCCTGCTAGTTCAGCGCTTTGTTCCCTCATCTTTTGTGTGGG 2186
Qy 2111 AGCTGATATCATGTCGATGCAACGCAATTTGATTAAGTGGAGAACTCTTAGTTTTTAT 2170
Db 2187 AGCTGATATCATGTCGATGCAACGCAATTTGATTAAGTGGAGAACTCTTAGTTTTTAT 2246
Qy 2171 TGCACAGAGGTCATTTTGGGGGCGTGCAACGCTGTTCTGGGTATTTCAATTTTTCATCAGCT 2230
Db 2247 TGCACAGAGGTCATTTTGGGGGCGTGCAACGCTGTTCTGGGTATTTCAATTTTTCATCAGCT 2306
Qy 2231 AGCATTTGATCTATGCTTTTGGGCTTCTCAGTTCATGAAAGCAATAATGAAGTATTT 2290
Db 2307 AGCATTTGATCTATGCTTTTGGGCTTCTCAGTTCATGAAAGCAATAATGAAGTATTT 2366
Qy 2291 AACTCTTTCATCAGTTCCTGCAAGTATGCTATTTAAATTTACTTTGGCCAGGATATAATG 2350
Db 2367 AACTCTTTCATCAGTTCCTGCAAGTATGCTATTTAAATTTACTTTGGCCAGGATATAATG 2426
Qy 2351 CCAGTCAGTCTCTTTATAGTGAAGAAATTTATTTGGTTAGTAATAATAATTTTAAACTA 2410
Db 2427 CCAGTCAGTCTCTTTATAGTGAAGAAATTTATTTGGTTAGTAATAATAATTTTAAACTA 2486
Qy 2411 AATATAAATCTATAATGTTTAAACATATGTTTCAATTAAGCAATGACATTTTGAATTA 2470
Db 2487 AATATAAATCTATAATGTTTAAACATATGTTTCAATTAAGCAATGACATTTTGAATTA 2546
Qy 2471 CTATATAATAGCTCATATTTACACTTACAGCTTTTCAATTTGATCAGGCTCTGAAATCTTT 2530
Db 2547 CTATATAATAGCTCATATTTACACTTACAGCTTTTCAATTTGATCAGGCTCTGAAATCTTT 2606

Qy	1352	ACCAACTCCAGTTGTTTAAAGCTCACTCATGTGACTGCACCTGAATCTAAGATCGGAAG	1411
Db	667	ACCAACTCCAGTTGTTTAAAGCTCACTCATGTGACTGCATCTGAATCTAAGGATCGGAAG	726
Qy	1412	AATCTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAATCAGCCATTTCCCTCTCTTAAGA	1471
Db	727	AATCTGTGGTTTCAGGAGGAAATTCCTTTTCAGAAATCAGCCATTTCCCTCTCTTAAGA	786
Qy	1472	TGGTTTTAGAGTCTTTGCCAGGAAGTGTAGGGTTTAACTTTGAAATAAAAATGGATCTGCC	1531
Db	787	TGGTTTTAGAGTCTTTGCCAGGAAGTGTAGGGTTTAACTTTGAAATAAAAATGGATCTGCC	846
Qy	1532	AGCAACGGGATGGAAATGTGGGATGGTAACTTATCAACATATTTTGAATGAAATCTGTTTT	1591
Db	847	AGCAACGGGATGGAAATGTGGGATGGTAACTTATCAACATATTTTGAATGAAATCTGTTTT	906
Qy	1592	TGGATATAATTTTAAAAAATGTTTTTGAATAATCTCGGAAGAGGAGAAATAGTGTTTCTT	1651
Db	907	TGGATATAATTTTAAAAAATGTTTTTGAATAATCTCGGAAGAGGAGAAATAGTGTTTCTT	966
Qy	1652	CATTGTGATGCAGATATTTGCAATGTTTCGGCAAAAAGCAGAAACAAATATCCGATACTAT	1711
Db	967	CATTGTGATGCAGATATTTGCAATGTTTCGGCAAAAAGCAGAAACAAATATCCGATACTAT	1026
Qy	1712	TTTTAACTCAAGGAAATCTGAGAAATTTATCCTGAATCTCATGGAACCTCAGATCTCGGACAA	1771
Db	1027	TTTTAACTCAAGGAAATCTGAGAAATTTATCCTGAATCTCATGGAACCTCAGATCTCGGACAA	1086
Qy	1772	CCCCATTGCAATGAGCTTTGCAAGTTTCGAAATCTACTGGGGATAAAATGTACATACTG	1831
Db	1087	CCCCATTGCAATGAGCTTTGCAAGTTTCGAAATCTACTGGGGATAAAATGTACATACTG	1146
Qy	1832	AAGACTTTGCTCAGAAACCCATCCTATATCAAGAGGCAAAAGCTTAAGGGACTAGTCAATAT	1891
Db	1147	AAGACTTTGCTCAGAAACCCATCCTATATCAAGAGGCAAAAGCTTAAGGGACTAGTCAATAT	1206
Qy	1892	TCTGCTGGGGTGATGATACCAATGATCCTGAAACAGAGGAAATTTGAAGCAACTTGGAG	1951
Db	1207	TCTGCTGGGGTGATGATACCAATGATCCTGAAACAGAGGAAATTTGAAGCAACTTGGAG	1266
Qy	1952	TTAATGCTCTAAATTTATGATAGGATATATGATTTGGATGCTTGAACCAACAAATATATTC	2011
Db	1267	TTAATGCTCTAAATTTATGATAGGATATATGATTTGGATGCTTGAACCAACAAATATATTC	1326
Qy	2012	AAGTGGAGCAATTTGGAACGGCTGAAGCAGGAAATTTGCCAGACTTTGTGTGTC	2071
Db	1327	AAGTGGAGCAATTTGGAACGGCTGAAGCAGGAAATTTGCCAGACTTTGTGTGTC	1386
Qy	2072	CCAATGTTAGCCGTTTGTTCCTCATCTTTGTGTGGGGATCTGATATCCATGTGGATG	2131
Db	1387	CCAATGTTAGCCGTTTGTTCCTCATCTTTGTGTGGGGATCTGATATCCATGTGGATG	1446
Qy	2132	CCAAAGGCAATTTGAACGCGAGAAATGCTTTAGTTTTTATTTGCAACAGAGTTCATTTTGGGG	2191
Db	1447	CCAAAGGCAATTTGAACGCGAGAAATGCTTTAGTTTTTATTTGCAACAGAGTTCATTTTGGGG	1506
Qy	2192	GCGTGCAACCGCTGTTCTGGGATATTCATTTTTCATCACTGAGCAATTTGTGATCTATGCCCT	2251
Db	1507	GCGTGCAACCGCTGTTCTGGGATATTCATTTTTCATCACTGAGCAATTTGTGATCTATGCCCT	1566
Qy	2252	TTGGGCTTCTCAGTTTCAATGAAGCAATTAATGAAGTATTTAACTCTTTTCACTACAGTTCTT	2311
Db	1567	TTGGGCTTCTCAGTTTCAATGAAGCAATTAATGAAGTATTTAACTCTTTTCACTACAGTTCTT	1626
Qy	2312	GCAAGTATGCTATTTAAATTTACTTGGCCAGGTATAATTTGCCAGTCACTCTTTTATAGTG	2371
Db	1627	GCAAGTATGCTATTTAAATTTACTTGGCCAGGTATAATTTGCCAGTCACTCTTTTATAGTG	1686
Qy	2372	AGAAAATTTTATGGTTAGTAAATATAAATATTTTAAACTAAATATAAATCTATAATGTT	2431
Db	1687	AGAAAATTTTATGGTTAGTAAATATAAATATTTTAACTAAATATAAATCTATAATGTT	1746

RESULT 8

RESOL 8
US-10-772-636-25

US-10-772-636-25
: Sequence 25. Application US/10772636

;; sequence 23, Application US/107
; Publication No. US20050042687A1

: GENERAL INFORMATION:

APPLICANT: Kelly, Louise M.

;
; APPLICANT: KELLY, LOUISE M.
;
; APPLICANT: CARROLL, JOSEPH M.

;
; AFFILIANT: CARROLL, JOSEPH
;
; APPLICANT: FARLOW, DEBORAH

; AFFILIANT: Farrow, Deborah
; APPLICANT: Healy, Aileen

1. TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,

21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,
; TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,

9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,
; TITLE OF INVENTION:

9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,
; TITLE OF INVENTION:

; TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR

; TITLE OF INVENTION: 5014

; FILE REFERENCE: MPI03-01SP1RNMNMIM

; CURRENT APPLICATION NUMBER: US/10/772,636

; CURRENT FILING DATE: 2004-02-05

; PRIOR APPLICATION NUMBER: US 60/445,241

;; PRIOR FILING DATE: 2003-02-05
;; PRIOR APPLICATION NUMBER: US 60/448,389
;; PRIOR FILING DATE: 2003-02-18
;; PRIOR APPLICATION NUMBER: US 60/456,320
;; PRIOR FILING DATE: 2003-03-20
;; PRIOR APPLICATION NUMBER: US 60/460,279
;; PRIOR FILING DATE: 2003-04-03
;; PRIOR APPLICATION NUMBER: US 60/465,924
;; PRIOR FILING DATE: 2003-04-28
;; PRIOR APPLICATION NUMBER: US 60/470,052
;; PRIOR FILING DATE: 2003-05-13
;; PRIOR APPLICATION NUMBER: US 60/498,106
;; PRIOR FILING DATE: 2003-08-26
;; PRIOR APPLICATION NUMBER: US 60/500,179
;; PRIOR FILING DATE: 2003-09-04
;; PRIOR APPLICATION NUMBER: US 60/502,909
;; PRIOR FILING DATE: 2003-09-15
;; PRIOR APPLICATION NUMBER: US 60/510,351
;; PRIOR FILING DATE: 2003-10-10
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 80
;; SOFTWARE: PaetSeq for Windows Version 4.0
;; SEQ ID NO 25
;; LENGTH: 2019
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(2019)
US-10-772-636-25

Query Match 63.0%; Score 2019; DB 21; Length 2019;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 145 ATGACACCTTCTCAGTTCGCTTGAATTAAGAGAACTCTTTTACAGGAGAGTCTTTT 204
DB 1 ATGACACCTTCTCAGTTCGCTTGAATTAAGAGAACTCTTTTACAGGAGAGTCTTTT 60
QY 205 GCGATATGGAAGCTGATGCTTTGGGAACTCGAATCCTCAAAATGCTGTGCTCTT 264
DB 61 GCGATATGGAAGCTGATGCTTTGGGAACTCGAATCCTCAAAATGCTGTGCTCTT 120
QY 265 CTTCCAGAGAAATGACACAGGTGAAGAGCTGATGGAAGCAACATTTGACTCAGTAGA 324
DB 121 CTTCCAGAGAAATGACACAGGTGAAGAGCTGATGGAAGCAACATTTGACTCAGTAGA 180
QY 325 GGAGTATCAGTTCAGTATCGCTACTTCAAGGGTACTTTTGAACCAAGACTATCGGT 384
DB 181 GGAGTATCAGTTCAGTATCGCTACTTCAAGGGTACTTTTGAACCAAGACTATCGGT 240
QY 385 GGTCCATGTCAGTATGATGATGTCAGAGTGGAGACTCATCTCAACCCAGATCAATAACC 444
DB 241 GGTCCATGTCAGTATGATGATGTCAGAGTGGAGACTCATCTCAACCCAGATCAATAACC 300
QY 445 CTTTGAAGCGAAATTTATTTAGCAGTGGACAAATTTGGAATCCACAATGGTGTGAA 504
DB 301 CTTTGAAGCGAAATTTATTTAGCAGTGGACAAATTTGGAATCCACAATGGTGTGAA 360
QY 505 ACTCTGGATTTCTGGATGGCTGACATGTCAGACTGGAATTAAGATTAGTTTGCATTTCT 564
DB 361 ACTCTGGATTTCTGGATGGCTGACATGTCAGACTGGAATTAAGATTAGTTTGCATTTCT 420
QY 565 GAAAAACCTCTGTGTCATTAACCAAGAAAAATTTAAAAATCTAGATTTAGGGTGAAG 624
DB 421 GAAAAACCTCTGTGTCATTAACCAAGAAAAATTTAAAAATCTAGATTTAGGGTGAAG 480
QY 625 CTGACACTAGAGCCCTGGAGGAGATGACGATGATAGGATCTCTCCACTGTACTCCAC 684
DB 481 CTGACACTAGAGCCCTGGAGGAGATGACGATGATAGGATCTCTCCACTGTACTCCAC 540
QY 685 AAAATGTCCAATAGCTTGGAGATATCCTTTAATAGCGCAATGAGTTCAAGTCAGGCAT 744

DB 541 AAAATGTCCAATAGCTTGGAGATATCCTTTAATAAGCGCAATGAGTTCAAGTCAGGCAT 600
QY 745 TCACAGCCGGAGTGTGGTTATGGCTTGCAGCCTGATCGTTGGACAGAGTACAGCATACAG 804
DB 601 TCACAGCCGGAGTGTGGTTATGGCTTGCAGCCTGATCGTTGGACAGAGTACAGCATACAG 660
QY 805 ACGATGGAAACAGATAAACCTGGAACATAATCTTTGATTTTTTTCGAAGAAGATCTCAGTGAG 864
DB 661 ACGATGGAAACAGATAAACCTGGAACATAATCTTTGATTTTTTTCGAAGAAGATCTCAGTGAG 720
QY 865 CACGTAGTTTCAAGGTGATGCCCTTCTCGACATGTGGGTACAGCTTGCTCTTATCATCC 924
DB 721 CACGTAGTTTCAAGGTGATGCCCTTCTCGACATGTGGGTACAGCTTGCTCTTATCATCC 780
QY 925 ACCATTGCTGAGAGTGGAAAGAGTCTCGAAATCTTACTCTTCCCATCATGAGCAAAAT 984
DB 781 ACCATTGCTGAGAGTGGAAAGAGTCTCGAAATCTTACTCTTCCCATCATGAGCAAAAT 840
QY 985 TCCCGGAAAAAATAAGGCAAGTGTGAGTGTGATATATAATTTAAGCCATTACAGGA 1044
DB 841 TCCCGGAAAAAATAAGGCAAGTGTGAGTGTGATATATAATTTAAGCCATTACAGGA 900
QY 1045 TACAGTTGTGACATGAATCTTCAATTTTCCAGTATTGGAAGCCCAAGATACCATTTGAT 1104
DB 901 TACAGTTGTGACATGAATCTTCAATTTTCCAGTATTGGAAGCCCAAGATACCATTTGAT 960
QY 1105 GTTGGCCATCGAGGTGCGAGAACTCTCAACAACTGCCCAGCTGGCTAAAGTTCAAGAA 1164
DB 961 GTTGGCCATCGAGGTGCGAGAACTCTCAACAACTGCCCAGCTGGCTAAAGTTCAAGAA 1020
QY 1165 AATACATTGCTTCTTTAAGAAATGCTGCTAGTCTAGTGTGAGCCCTTTAGAAATTTGAC 1224
DB 1021 AATACATTGCTTCTTTAAGAAATGCTGCTAGTCTAGTGTGAGCCCTTTAGAAATTTGAC 1080
QY 1225 GTAACACCTTTCAAGAGCTTTGTGTCCTGCTGATATATCATGATCTTACCTGTTTGTGACT 1284
DB 1081 GTAACACCTTTCAAGAGCTTTGTGTCCTGCTGATATATCATGATCTTACCTGTTTGTGACT 1140
QY 1285 ATGAAAAGAAATTTGATGCTGATCCAGTGTGAATTTTGAATTTCCAGTAAAGAAATTA 1344
DB 1141 ATGAAAAGAAATTTGATGCTGATCCAGTGTGAATTTTGAATTTCCAGTAAAGAAATTA 1200
QY 1345 ACATTTGACCAACTCCAGTGTGTTAAAGCTCACCTCATGTGACCTGCACTGAAATCTAAGGAT 1404
DB 1201 ACATTTGACCAACTCCAGTGTGTTAAAGCTCACCTCATGTGACCTGCACTGAAATCTAAGGAT 1260
QY 1405 CGGAAAGAAATCTGTGTTTCAAGAGGAAATTCCTTTTCAGAAATCAGCCATTTCTTCT 1464
DB 1261 CGGAAAGAAATCTGTGTTTCAAGAGGAAATTCCTTTTCAGAAATCAGCCATTTCTTCT 1320
QY 1465 CTTAAGATGTTTTAGAGTCTTTGCCAGAGATGTAGGGTTTAAACATTTGAAAATAAATGG 1524
DB 1321 CTTAAGATGTTTTAGAGTCTTTGCCAGAGATGTAGGGTTTAAACATTTAAGATTAAGTGG 1380
QY 1525 ATCTGCCAGCAAGGGATGGAATGTGGGATGGAATCTTATCAACATATTTTGAATGAATG 1584
DB 1381 ATCTGCCAGCAAGGGATGGAATGTGGGATGGAATCTTATCAACATATTTTGAATGAATG 1440
QY 1585 CTGTTTTTGGATATAATTTTAAAAATCTGTTTTAGAAAATCTGCGGAGAGAGAGATAGTG 1644
DB 1441 CTGTTTTTGGATATAATTTTAAAAATCTGTTTTAGAAAATCTGCGGAGAGAGAGATAGTG 1500
QY 1645 TTTTCTTCATTTTGTATGAGATATTTTGCACAAATGTTTGGCAAGAGAGAGAGAGATATCCG 1704
DB 1501 TTTTCTTCATTTTGTATGAGATATTTTGCACAAATGTTTGGCAAGAGAGAGAGAGATATCCG 1560
QY 1705 ATACTATTTTAACTCAAGGAAATCTGAGATTTTATCTGAACCTCATGAGCCTCAGATCT 1764
DB 1561 ATACTATTTTAACTCAAGGAAATCTGAGATTTTATCTGAACCTCATGAGCCTCAGATCT 1620
QY 1765 CGGACAAACCCCATTTGCAATGAGCTTTGCAAGTTTGAATTTTGAATTTTGAATTTGAATGTA 1824
DB 1621 CGGACAAACCCCATTTGCAATGAGCTTTGCAAGTTTGAATTTTGAATTTTGAATTTGAATGTA 1680

Db 988 CGCTGGAACTCTTACTTCCCATCATGACGAGAGTTCAGAAAAAATATAGGCAAG 1047
Qy 1007 TGAGATTGACTATATATATTAAGCATTAACAGGATACAGTTGACATGAATCTT 1066
Db 1048 TCAGATTGATTTTATCATCATCAAGCAATACCCAGGATATAGTTGTTTATGACGCTT 1107
Qy 1067 CATTTTCAAGTATTTGGAGCAAGATACCAATTTGGATTTGGCCATCGAGGTGCAGAA 1126
Db 1108 CATTTCCAACTATTTGGAACCAAGATACCACTGGATTTGGACATCGTGGTGCAGGA 1167
Qy 1127 ACTCTACAACTGCCAGCTGGCTAAAGTTCAAGAAAAATATATTTGCTTTTAAAGAA 1186
Db 1168 ACTCAACAACTGCCAAGCTGGCTAAAGTACAGGAATATATTTGCTTTTAAAGAA 1227
Qy 1187 ATGCTGCTAGTCACTGGTGCAGCTTTTGPAGATTTGACGTACACCTTTCAAAGGACTTTG 1246
Db 1228 ATGCTGCCAGCAGTGGTGCAGCAATTTTGGAAATTTGATGTCCACCTTTCAAAGGACTTAG 1287
Qy 1247 TGCCCGTGGTATATCATGATCTTACCTGTTGTTGACTATCAAAAAGAAATTTGATGCTG 1306
Db 1288 TGCTGTAGTATCATGATCTCACCTGCTGTTAACTATGAAAAGGAAATATGAAGCTG 1347
Qy 1307 ATCCAGTTGAATTTTCAAAATTCAGTAAAGAAATTAACATTTGACCACTCCAGTTGT 1366
Db 1348 ATCCAGTTGAATTTTGAATTCACGTAAGAAATTAACATTTGACCACTCCAGTTAT 1407
Qy 1367 TAAAGCTCACTCATGCTGACCTGCACTGAAATTAAGGATCGGAAGAAATCTCTGTTGTCAGG 1426
Db 1408 TGAAGCTTTCATGCTGACTGCACTAAAAACCAAGACCAAGAAACAAATGTATGCTGAGG 1467
Qy 1427 AGGAAATTCCTTTTCAGAAATTCAGCCATTCCTTCTCTTAAGATGTTTTAGAGCTT 1486
Db 1468 AGGAAATTCCTTTTCAGAAATTCAGCCATTCCTTCTCTTAAGATGTTTTAGAGCTAT 1527
Qy 1487 TGCCAGAACTGAGTTTACATTTGAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1546
Db 1528 TGCCAGAAATGAGGATTTATATAGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1587
Qy 1547 TGTGGATGGTAACTTATCAACATATTTTGAATGAAATGAAATGAAATGAAATGAAATGAA 1606
Db 1588 TATGGAGCGCACTTATCGACATATTTTGAATGAAATGAAATGAAATGAAATGAAATGAA 1647
Qy 1607 AAATCTTTTGAAGATTTGGAAGAGAGAGATAGTTTCTTCTTCAATTTGATGAGATA 1666
Db 1648 AAATCTTTTGAAGATTTGGAAGAGAGAGATAGTTTCTTCTTCAATTTGATGAGATA 1707
Qy 1667 TTTGCAATGTTTCGGCAAGACGACAAATATCCGATCTATTTTAACTCAAGGAA 1726
Db 1708 TCTGTACAATGTTTCGGCAAGACGACAAATATCCGATCTATTTTAACTCAAGGAA 1767
Qy 1727 AATCTCAGATTTTATCTGAACTCATGACCTCAGATCTCGGACAAACCCCAATGCAATGA 1786
Db 1768 AGTCTGACATTTTACCTTGAACTCATGACCTCAGATCTCGGACAAACCCCAATGCAATGA 1827
Qy 1787 GCTTTGCACAGTTTGAAGATTTTCTACTGGGATAAATGTAATATGAGACTTGTCTCAGAA 1846
Db 1828 GCTTTGCACAGTTTGAAGATTTTCTACTGGGATAAATGTAATATGAGACTTGTCTCAGAA 1887
Qy 1847 ACCATCTCTATTTCAAGAGCAAGACTTAAGGACTAGTCTATTTCTCTGCTGGGGTATG 1906
Db 1888 ACCATCTCTATTTCAAGAGCAAGACTTAAGGACTAGTCTATTTCTCTGCTGGGGTATG 1947
Qy 1907 ATACCATGATCTCTGAAACAGAGGAAATTTGAAGGAACTTGGAGTTAATGTTCTAATTT 1966
Db 1948 ATACCATGATCTCTGAAACAGAGGAACTTGGAGTTAATGTTCTAATTTGTTCTAATAT 2007
Qy 1967 ATGATAGG 1974
Db 2008 ATGATAGG 2015

; Sequence 48, Application US/10426776
; Publication No. US20040009553A1
; GENERAL INFORMATION:
; APPLICANT: Gluckesman, Maria Alexandra
; APPLICANT: Williamson, Mark J.
; APPLICANT: Tsia, Fong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Kapeller-Ibermann, Rosana
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Chiang, Lillian Wei-Ming
; APPLICANT: Hunter, John Joseph
; APPLICANT: Wood, Andrew
; APPLICANT: Jenkins, Lorayne P.
; TITLE OF INVENTION: NOVEL 27411, 23413, 22438, 23553,
; TITLE OF INVENTION: 25278, 26212, NARC SCI, NARC 10A, NARC 1, NARC 12, NARC 13,
; TITLE OF INVENTION: NARC17, NARC 25, NARC 3, NARC 4, NARC 7, NARC 8, NARC 11,
; TITLE OF INVENTION: NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20, NARC 26, NARC
; TITLE OF INVENTION: 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C, NARC
; TITLE OF INVENTION: 8B, NARC 9, NARC2A, NARC 16B, NARC 1C, NARC 1A, NARC 25,
; FILE REFERENCE: MPI03-0620NMIM
; CURRENT APPLICATION NUMBER: US/10/426,776
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 10/229,662
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/795,691
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,517
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 10/105,992
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/406,045
; PRIOR FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 10/314,881
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 09/773,426
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 09/495,823
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/692,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/161,188
; PRIOR FILING DATE: 1999-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 2393
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-10-426-776-48

Query Match 30.9%; Score 991; DB 17; Length 2393;
Best Local Similarity 80.1%; Pred. No. 1.1e-217;
Matches 1418; Conservative 0; Mismatches 265; Indels 88; Gaps 18;

Qy 1284 TATCAAAAAGAAATTCATGCTGATCCAGTTGAAATTTTGAATTCAGTAAAGAAATTCAGTAAAGAAATTC 1343
Db 61 TTTCTAATAGAAATATGAAGCTGATCCAGTTGAAATTTGTTTGAATTCAGTAAAGAAATTCAGTAAAGAAATTC 120
Qy 1344 AACATTTGACCAATCCAGTTGTTTAAAGCTCACTCATGTGACCTGACCTGAAATCTAAGGA 1403
Db 121 AACATTCGACCAATCCAGTTTATGAAGCTTCTCATGTGACCTGACCTAATAAACCAAGA 180
Qy 1404 TCGGAAAGAAATCTGTGTTTCAAGGAGAAATTCCTTTTCAGAAATCAGCCATTTCTTTC 1463
Db 181 CCAGAAACAAATGATGCTGAGGAGGAAATTCCTTTTCTGAAACCAACCATTTCTTCTTC 240
Qy 1464 TCTTAAGATGTTTTAGAGTCTTTTCCAGAGAGATGTTAGGGTTTAAACATTTGAAATATAAATG 1523
Db 241 TCTTAAGATGTTTTAGAGTCAITTCAGGAGGAAATTTAGGATTTTAAATATAGAAATATAAATG 300
Qy 1524 GATCTGCCAGCAAGGGGATGGAATGATGGGATGGAATCAATTAATCAATATTTTTCATGAA 1583

Db	301	GATTTGCCAACACAGGGATGAGATGGGACGGCAACTTATCGACATATTTTGGATGAA	360
Qy	1584	TCGTGTTTTGGATATAATTTTAAAAAAGCTGTTTTAGAAAAATCTCTGGGAAGAGGAGAATAGT	1643
Db	361	TGCATTTTGGATATAATTTTAAAAAAGCTGTTTTAGAAAAATCTGGGAAGAGGAGAATAGT	420
Qy	1644	GTTTTCTTCATTTTGATGCAGATATTTGCACAATGGTTTCGGCAAAAGCAGAACCAATATCC	1703
Db	421	ATTTTCTTCATTTTGATGCAGACATCTGTACAATGGTTTCGGCAGAAAACAAAACCAATATCC	480
Qy	1704	GATACATATTTTAACTCAAGGAAATCTGAGATTTTATCCTGAACTCATGACCTCAGATC	1763
Db	481	CATATTTATTTTGAACCAAGGAAAGCTGACATTTTACCTCTGAACTCATGACCTCAGATC	540
Qy	1764	TCGGACAACCCCAATTCGAATGAGCTTTGCACAGTTTGGAAAAATCTACTCTGGGGAATAATGT	1823
Db	541	TCGGACAACACCATTGCATGAGCTTTGCACAGTTTGGAAAAATATTTTGGGGAATAATGC	600
Qy	1824	ACATACTGAAGACTTGTCTCAAAACCCATCCTATATTTCAAGAGGCAAAAGCTAAGGSACT	1883
Db	601	CCATCTGAAGATCTCTTTAGAAACCCATCCTATGTCTCAAGAGGCAAAAGATAAGGSAAT	660
Qy	1884	AGTCATATTTCTGCTGGGGTGATGATACCAATGATCTGTGAAAAACAGAAAGGAAATTTGAAGGA	1943
Db	661	GGTCATATTTCTGCTGGGGTGATGATACCAATGATCTGTGAAAAACAGAAAGGAAATTTGAAGGA	720
Qy	1944	ACTTTGGAGTTAATGGTCTAAATTTATGATAGATATATGATTTGGATGCTCTGAAACCAAA	2003
Db	721	ATTTGGAGTAAATGGTCTAAATATATGATAGATATGATGATTTGGATGCTCTGAAACCAAA	780
Qy	2004	TATATTTCAAAGTGGAGCAAATGGAAACCCCTGAAGACAGAAATTCGCAGAGCTTAAAGAGCTG	2063
Db	781	TATATTTCAAAGTGGAGCAAATTCGAGGCGCTGAAGCGAGAAATGCCAGAGCTTAAAGAACTG	840
Qy	2064	TTTGTGTCCCACTGTTAGCGCGTTTGTTCCTCATCTTTGTGTGGGGAGTCTGATATCCA	2123
Db	841	TTTGTGTGCCCACTGTTAGCCACTTTCATTCCTCTCTCTCTCTGTA.TGGAGTCTAAAAATCCA	900
Qy	2124	TGTGGATGCCAACCGGCAATGCATACGTGGAGAAATGCTTAGTTTTATTGACACAGAGGTCA	2183
Db	901	TGTGGATGCTAACCGCAATGATATGTGGAGAACGCTTAGTTCCTAGTGCACAGAGGACA	960
Qy	2184	TTTTGGGGGCGTGCACGCGTGTCTGGGTATTCATTTTTTCATCACTGAGCATTTGTTGATC	2243
Db	961	TTCAAGAGGCTCTCCCTCGCGCTGAG---TTCGTCTCCACCACTGAAACACCGGTC---	1013
Qy	2244	TATGCTTTTGGGGTCTCTAGTTCATGAAGCAATAATGAAGTATTTAACTCTTTCACTA	2303
Db	1014	---GCCTCTTTAGGTTCTCAGTCCAAATGAAGCAATATGAAGTATTTTAA---CTATCATTTA	1068
Qy	2304	CAGTCTTTCGAAGTA-----TGCTATTTAAATTACTTCGCCCAGGTATAAATTCGCCAGTC	2356
Db	1069	CAGTTCGCCGAAGAAATATCAAGTACACTATTTATCACTGTCTCAGGTATTAATTTACCAATC	1128
Qy	2357	AGTCTCTTTATAGTGAGAAAAATTTATTTGGTTAGTAATATAAATATTTTAACTTAAATATA	2416
Db	1129	AGTCTCTGTACA-----AA.TGTTAAACACTTTAATGAGAGATC	1167
Qy	2417	TAAATCTATAATGTTTAAACATATGTTTCATTTAAAGCATAGCACTTTTGAATTTAACTATAT	2476
Db	1168	TAAAGCCTATAATGGTGA---ATCTTTCATTTAAAGCATATACTTTGGAATTAAGCTATAT	1223
Qy	2477	AAATAGCTCATATTTACACTTACAGCTTTTCATTTGATCAGTCTGAAATCTTTAGCACT	2536
Db	1224	AAATATCTCAT-----AGTTCAGGCTTTTCATTTTGAATTAGGTCCTTAAATCTTCAGTGCT	1277
Qy	2537	TAAAGGAAATGACTATGCATAAATTTATACCTGCAGCCATGAAAAAATAAGTACCTCAATGC	2596
Db	1278	TGAGAAAAATGAC--TGCATAAATTTACCTGCACCATGGAATTAATAGTACCTCAAGTGC	1335
Qy	2597	ATGCATTTGCATCTGGTGATTTCCAACTGCACAAATCTTTGTGCCATCTTGTATATAGGTAT	2656

1336	Db	ATGCATTGTCACCTGGTGGCTCCAGCTGCACAAAGCTGTGTGT-CATCCATGTTACATAGGTGT	1394
2657	Qy	TTTTTACATGGGTGGACATGCACACAAACCACTTTTTCATTTCAGTATGAACCTTTGAGGCTG	2716
1395	Db	C-TTTACATGGGTGGATAGAAAATG-CACTAGGCTCCTTTAGTATAAACCTC-AGACTG	1451
2717	Qy	CTGCATTTTTCCACTTTAAACCAACAGAGCTGAAGGTGAACCTTCGAAACTTTGTTTCATAA	2776
1452	Db	CTGCATTTTCCCACTGACCCAAACCAAGCTCGAGATGAACCTCAAAACTTTGTTTCATAG	1511
2777	Qy	ATCTTTCAAAAGTTGTTTTTACATCAATGTTTAAAAATTTCAAAATGCTGCAGGGTAAATTTAA	2836
1512	Db	ACTGTTCAAA-----GATTTTAAAGTTCAGAAATGCTGCAGGGTAACTTTAA	1558
2837	Qy	TGTATAAAATPATTTAGTAAGAAAAAGTATGTTATTCATACCTAGTAGAATAGATCAACA	2896
1559	Db	TGTATAAAGTATTTGTAAG---AGGTATATATTGTCATATATAGTCGTGTAGATCAGATG	1615
2897	Qy	TACAAATTCAAATTCAGTCGATGCTTTAG----GTGTTAAGCATGAGATTGTACATGTTTA	2952
1616	Db	TGTAATTTTGACTCCGTGCGATGCTTTAGGTTTGTTTTAAGCATGAGGTTGTACATGTTTA	1675
2953	Qy	CTGTTTAGGTCCTTGCATCTGTGTGTGTCTAGGTAGTATGAAGAATGTCGAAGCTGGACG	3012
1676	Db	CT-----GTCCCTTCATC--TGGTGTAGGTGAGT-----GAGATGTTAAGGACTGAAAA	1723
3013	Qy	TATTTTGTGGCTTAAAAAAAAGGCTGTTT	3043
1724	Db	TATTTTGTCCCTTAAAAAGAGTATGCCTATT	1754

RESULT 11

```

US-09-796-692-4214
; Sequence 4214, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND
; TITLE OF INVENTION: HEMATOLOGICAL MA
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version
; SEQ ID NO 4214
; LENGTH: 524
; TYPE: DNA

```

```
; ORGANISM: Homo sapiens
US-09-796-692-4214
Query Match      16.3%; Score 522.4; DB 9; Length 524;
Best Local Similarity 99.8%; Pred. No. 6.1e-110;
Matches 523; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2405 AAACCTAAATATATAAATCTAATGTTAAACATATGTTCAATTAAGCATAGCACTTTGA 2464
Db 1 AAACCTAAATATATAAATCTAATGTTAAACATATGTTCAATTAAGCATAGCACTTTGA 60
QY 2465 AATTAACCTATATAATAGCTCATATTTACACTTACAGCTTTTCATTGATGAGTCTGAA 2524
Db 61 AATTAACCTATATAAATGCTCATATTTACACTTACAGCTTTTCATTGATGAGTCTGAA 120
QY 2525 ATCTTTAGCACTTAAGGAAATGACTATGCAATAATTAACCTGACCAATGAAAAAATAAG 2584
Db 121 ATCTTTAGCACTTAAGGAAATGACTATGCAATAATTAACCTGACCAATGAAAAAATAAG 180
QY 2585 TACCTCAAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2644
Db 181 TACCTCAAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 2645 GTATATAGGTATTTTACATGGGTTGACATGCGACACAAACCACTTTTCATTTCAGTATGA 2704
Db 241 GTATATAGGTATTTTACATGGGTTGACATGCGACACAAACCACTTTTCATTTCAGTATGA 300
QY 2705 ACCTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2764
Db 301 ACCTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 2765 CTGTGTTTCATAAATCTTTCAAAAGTTGTTTACATCAATGTTAAATTTTCAAAATGCTGC 2824
Db 361 CTGTGTTTCATAAATCTTTCAAAAGTTGTTTACATCAATGTTAAATTTTCAAAATGCTGC 420
QY 2825 AGGTAATTTAATGATATAAATATTAGTAAGAAAGATGATGATGCACTATTAGTAGAA 2884
Db 421 AGGTAATTTAATGATATAAATATTAGTAAGAAAGATGATGATGCACTATTAGTAGAA 480
QY 2885 TAGTCACAAACATACAAATTCATTCAGTGCATGCTTTAGGTGT 2928
Db 481 TAGTCACAAACATACAAATTCATTCAGTGCATGCTTTAGGTGT 524

RESULT 12
US-10-040-862-4214
; Sequence 4214, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
```

```
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4214
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-4214
Query Match      16.3%; Score 522.4; DB 14; Length 524;
Best Local Similarity 99.8%; Pred. No. 6.1e-110;
Matches 523; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2405 AAACCTAAATATATAAATCTAATGTTAAACATATGTTCAATTAAGCATAGCACTTTGA 2464
Db 1 AAACCTAAATATATAAATCTAATGTTAAACATATGTTCAATTAAGCATAGCACTTTGA 60
QY 2465 AATTAACCTATATAATAGCTCATATTTACACTTACAGCTTTTCATTGATGAGTCTGAA 2524
Db 61 AATTAACCTATATAAATGCTCATATTTACACTTACAGCTTTTCATTGATGAGTCTGAA 120
QY 2525 ATCTTTAGCACTTAAGGAAATGACTATGCAATAATTAACCTGACCAATGAAAAAATAAG 2584
Db 121 ATCTTTAGCACTTAAGGAAATGACTATGCAATAATTAACCTGACCAATGAAAAAATAAG 180
QY 2585 TACCTCAAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2644
Db 181 TACCTCAAAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 2645 GTATATAGGTATTTTACATGGGTTGACATGCGACACAAACCACTTTTCATTTCAGTATGA 2704
Db 241 GTATATAGGTATTTTACATGGGTTGACATGCGACACAAACCACTTTTCATTTCAGTATGA 300
QY 2705 ACCTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2764
Db 301 ACCTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 2765 CTGTGTTTCATAAATCTTTCAAAAGTTGTTTACATCAATGTTAAATTTTCAAAATGCTGC 2824
Db 361 CTGTGTTTCATAAATCTTTCAAAAGTTGTTTACATCAATGTTAAATTTTCAAAATGCTGC 420
QY 2825 AGGTAATTTAATGATATAAATATTAGTAAGAAAGATGATGATGCACTATTAGTAGAA 2884
Db 421 AGGTAATTTAATGATATAAATATTAGTAAGAAAGATGATGATGCACTATTAGTAGAA 480
QY 2885 TAGTCACAAACATACAAATTCATTCAGTGCATGCTTTAGGTGT 2928
Db 481 TAGTCACAAACATACAAATTCATTCAGTGCATGCTTTAGGTGT 524

RESULT 13
US-10-057-475B-4214
; Sequence 4214, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
```


Db 301 ACCTTGAGCTGCTGCCATTTTCCACTTAACCAACGAGCTGAAGCTCGAAA 360
QY 2765 CTTGTTTCATAAACTTTTCAAAGTGTGTTTACATCAATGTTAAATTTCAAATGCTGC 2824
Db 361 CTTGTTTCATAAACTTTCAAAGTGTGTTTACATCAATGTTAAATTTCAAATGCTGC 420
QY 2825 AGGTPAATTAATGATATAAATATAGTAAGAAAAGTATGTTGATACCTTAGTAGAA 2884
Db 421 AGGTPAATTAATGATATAAATATAGTAAGAAAAGTATGTTGATACCTTAGTAGAA 480
QY 2885 TAGATCAACATACAAATTCATTCAGTGCATGCTTTAGGTGT 2928
Db 481 TAGATCAACATACAAATTCATTCAGTGCATGCTTTAGGTGT 524

RESULT 15
US-10-764-324-4214
; Sequence 4214, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4214
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-4214

Query Match 16.3%; Score 522.4; DB 19; Length 524;
Best Local Similarity 99.8%; Pred. No. 6.1e-110;
Matches 523; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2405 AAACATAATATAAATCTATAATGTTAAACATATGTTTAAAGAGCATAGCATTGGA 2464
Db 1 AAACATAATATAAATCTATAATGTTAAACATATGTTTAAAGAGCATAGCATTGGA 60
QY 2465 AATTAACATATAAATAGCTCATATTTACATTTACAGCTTTTCATTTGATCAGGTCTGAA 2524
Db 61 AATTAACATATAAATAGCTCATATTTACATTTACAGCTTTTCATTTGATCAGGTCTGAA 120
QY 2525 ATCTTTAGCATTAAAGAAAATGACTATGATGATAATTATACCTGACCATGAAAAAATAAG 2584

Db 121 ATCTTTAGCATTAAAGAAAATGACTATGCTATAATTATACCTGACCATGAAAAAATAAG 180
QY 2585 TACCTCAAAATGCATGCAATTTGCACTGGTGATTCCAACTGCACAAAATCTTTTGCCATCTT 2644
Db 181 TACCTCAAAATGCATGCAATTTGCACTGGTGATTCTAACTGCACAAAATCTTTTGCCATCTT 240
QY 2645 GTATATAGGTATTTTATACATGGGTTGACATGCAACAACACCAATTTTCATTCAGTATGA 2704
Db 241 GTATATAGGTATTTTATACATGGGTTGACATGCAACAACACCAATTTTCATTCAGTATGA 300
QY 2705 ACCTTGAGGCTGCTGCCATTTTCCATTAACCAAAACGAGCTGAAGGTGAACCTCGAAA 2764
Db 301 ACCTTGAGGCTGCTGCCATTTTCCATTAACCAAAACGAGCTGAAGGTGAACCTCGAAA 360
QY 2765 CTTGTTTCATAAATCTTTCAAAGTGTGTTTACATCAATGTTTAAATTTCAAATGCTGC 2824
Db 361 CTTGTTTCATAAATCTTTCAAAGTGTGTTTACATCAATGTTTAAATTTCAAATGCTGC 420
QY 2825 AGGTAATTTAATGTATAAAATATTAGTAAGAAAAAGTATGTTTGCATACCTTAGTAGAA 2884
Db 421 AGGTAATTTAATGTATAAAATATTAGTAAGAAAAAGTATGTTTGCATACCTTAGTAGAA 480
QY 2885 TAGATCAACATACAAATTCATTCAGTGCATGCTTTAGGTGT 2928
Db 481 TAGATCAACATACAAATTCATTCAGTGCATGCTTTAGGTGT 524

Search completed: July 3, 2005, 00:06:53
Job time : 1867 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 3, 2005, 06:48:25 ; Search time 5762 Seconds

(without alignments)
4439.290 Million cell updates/sec

Title: US-10-047-855-3

Perfect score: 3537

Sequence: 1 MTPQVAFIRGTLLPGEVF.....LCGSIDIHVANGINDVNA 672

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10047855/runat_01072005_155529_2232/app_query.fasta_1.839
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10047855@cgn 1.1 4352 @runat_01072005_155529_2232 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gssi:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3331.5	94.2	3142	3 AK030645	Mus muscu
2	3318.5	93.8	3614	3 AK050318	Mus muscu
3	3150	89.1	3248	3 BC006887	Mus muscu
4	3066	86.7	3578	3 AK031658	Mus muscu
5	2400	67.9	2604	3 AK009563	Mus muscu
6	2143	60.6	1618	3 AK049491	Mus muscu
7	2119	59.9	2228	3 AK051728	Mus muscu
8	1570	44.4	1006	5 BX344123	BX344123
9	1486.5	42.0	1128	4 BM478430	BM478430 AGENCOURT

10	1365.5	38.6	901	6 CB196769	CB196769 AGENCOURT
11	1322	37.4	882	4 BM457674	BM457674 AGENCOURT
12	1321	37.3	818	5 BX403389	BX403389
13	1319.5	37.3	783	6 CD466056	CD466056 Leukon2 1
14	1305	36.9	958	5 BU503488	BU503488 AGENCOURT
15	1285	36.3	948	5 BU145581	BU145581 AGENCOURT
16	1260	35.6	754	7 CN303372	CN303372 170006000
17	1248	35.3	833	4 BU144681	BU144681 602909978
18	1227.5	34.7	855	6 BQ962080	BQ962080 AGENCOURT
19	1223.5	34.6	805	6 CA317224	CA317224 UI-N-FW0-
20	1218.5	34.5	878	5 BQ930039	BQ930039 AGENCOURT
21	1216.5	34.4	728	7 CN303370	CN303370 170004243
22	1216	34.4	696	2 AW327746	AW327746 dr01904.x
23	1186	33.5	688	4 BM786845	BM786845 K-EST0065
24	1185	33.5	711	5 BX507382	BX507382 DKF2p779K
25	1156.5	32.7	691	7 CK833283	CK833283 4057562 B
26	1153.5	32.6	821	4 BU822275	BU822275 603036411
27	1148.5	32.5	836	5 BU519577	BU519577 AGENCOURT
28	1135	32.1	770	6 CB236281	CB236281 AGENCOURT
29	1132	32.0	786	7 CO795305	CO795305 AGENCOURT
30	1104	31.2	719	7 CO433958	CO433958 UI-N-HX0-
31	1103	31.2	660	6 BY709247	BY709247 BY709247
32	1102.5	31.2	768	1 AU136888	AU136888 AU136888
33	1095.5	31.0	795	4 BG623187	BG623187 60248092
34	1089	30.8	667	7 CK832619	CK832619 4056677 B
35	1075	30.4	665	6 BY734861	BY734861 BY734861
36	1066.5	30.2	760	7 CF286044	CF286044 AGENCOURT
37	1056	29.9	748	7 CO559324	CO559324 AGENCOURT
38	1041	29.4	667	6 CD771503	CD771503 AGENCOURT
39	1032.5	29.2	665	6 CD640091	CD640091 AGENCOURT
40	1030	29.1	764	7 CV073042	CV073042 AGENCOURT
41	1015.5	28.7	621	7 CK623779	CK623779 mi13e07.y
42	1007	28.5	642	2 BB652394	BB652394 BB652394
43	994.5	28.1	746	7 CN303375	CN303375 170005318
44	994	28.1	578	7 CV024277	CV024277 1525 Full
45	986	27.9	861	7 CO554090	CO554090 AGENCOURT

ALIGNMENTS

RESULT 1
AK030645 3142 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430401018 product:hypothetical Glycero-phosphoryl diester phosphodiesterase/Glycosyl hydrolase, starch-binding domain containing protein, full insert sequence.
DEFINITION AK030645 GI:26326638
ACCESSION AK030645.1
VERSION HTG; CAP trapper.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Normalizaton and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3142)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
Location/Qualifiers
1. .3142
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:5430401018"
/db_xref="taxon:10090"
/clone="5430401018"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="6 days neonate"
157..2184
/note="unnamed protein product; hypothetical
Glycerophosphoryl diester phosphodiesterase/glycosyl
hydrolase, starch-binding domain containing protein
(InterPro|IPR004129, InterPro|IPR002044, evidence:
InterPro)
putative"
/codon_start=1
/protein_id="BAC27063.1"
/db_xref="GI:26326639"
/translation="MTSPQVTFEIRGTLGPGEVFAICGSCDALGNWNPQNAVALINEN
ETGDSVLKAVIALNRGVSKYFRPGCFLEPKTIGGPGQVIVHKWETHLQPSITPL
ESEIIDDGGQFHNGVETLDSGLTCTQTEIRLKLHFSEKPPVSIKKKFKRFRVK
LITLGLGDEDDDDKVSPTVILHKWNSLEISLISDNFEKRSQPCGCGLOPDRWT

EYSIOTMEPNLELIPDFFEEDLSEHVQDVLPGHVGTACLSSSTIABSGRSAGILT
LPTMSNRKTIGKVRVDLIIKPLPGYSCMOSSFSKYWKPRIPLDVGHRGAGNSTT
TAKLAKVQENTIASLNRASAGAAFEVFDVHLSKDFVUVYHDLTCLTWMKRYEADP
VELUFEIPELTFDQQLKSLTAKDKQSLYEENFESENPQFSLKMWLES
LPENVGFIIEIKWICQHRDGVMDNLTSTYFDMNVFLDIILKTVLENSGKRRIVFSFSD
ADICTMVRQNKYPIFLFTQKSDITYPELMDLRSRTTIAMSFQAFENILGINAHT
DILLNPSYVQEAQAKGLVIFCWGDDTNDPENRRKLKEFGVNGLIYDRIYDMPEQPN
FOVEQLERLKLQELPELKNCLCPTVSHFIPSSFCVEPDIIHVDANGIDSVENA"
3125..3130
/note="putative"
3142
/note="putative"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3142
Score: 3331.50 Matches: 624
Percent Similarity: 97.04% Conservatives: 31
Best Local Similarity: 92.44% Mismatches: 17
Query Match: 94.19% Indels: 3
DB: 3 Gaps: 1
US-10-047-855-3 (1-672) x AK030645 (1-3142)
QY 1 MetThrProSerGlnValAlaPheGluLeuArgGlyThrLeuLeuProGlyGluValPhe 20
Db 157 ATGACACCTTCTCAGGTCACTTTTGAAATAAGAGGAACCTCTTTTACCAGGAGGCTCTT 216
QY 21 AlaileCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 217 GCATATGTGGAGCTGTGATGCGCTGGGAACTGGAACTCTCAAAATGCTGTGGCTCTT 276
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 277 ATTAATGAAACAGAGAGAGAGAGAGTGTCTTGTGAAAGCAGTGTCTCTCAATAGA 336
QY 61 GlyValSerValGlnTrpArgTrpPheLysGlyTyPheLeuGluProLysThrIleGly 80
Db 337 GGAGTGTGAGTGAAGTACCGCTACTTTCAGAGGCTGCTTTTGTAGAACCAAGATATCGGT 396
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 397 GGTCCATGTCAGTCACTAGTTCACAGTGGAGACTCATCTACACCAACCATCAATACC 456
QY 101 ProLeuGluSerGluIleLeuIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
Db 457 CCTTTAGAAAGTCAAATCATTATTGACATGGACAGTTTGGCATCCACAATCGTGTGAA 516
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTySer 140
Db 517 ACACCTGGATTCTGGATGGCTTACATGTCAGACTGAAATAGATGGCTCTGCCATTTTCT 576
QY 141 GluLysProValSerIleThrLysLysLeuLysLysSerArgPheAspValLys 160
Db 577 GAGAAACCTCTCTGTTCAATTACCAAGAAAGAGTTCAAAAATCGAGATTAGGTTAAG 636
QY 161 LeuThrLeuGluGlyLeuGluGlu-----AspAspAspAspArgValSerProThr 177
Db 637 CTCACACTCGAGGCTCTGGAGGAAGATGAAGATGATGATGATGATGATGATGATGATGAT 696
QY 178 ValLeuHisLysMetSerAsnSerLeuGluLysSerLeuLysSerLeuLysSerLeuLys 197
Db 697 GTTCTTCAAAAATGTCACACAGCTGGAGATATCCTTAATAAGTACATGATGATGATGATGAT 756
QY 198 CysArgHisSerGlnProGluCysGlyTyGlyLeuGlnProAspArgTrpThrGluTy 217
Db 757 TGCAGGACTCTCAGCCAGCAATGTGGGTATGGCTTACAGCCCATGATGATGATGATGATGAT 816
QY 218 SerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPhePheGluGluAsp 237
Db 817 AGCATACAGACATGAAACCAAGATATCTGGAGCTCATCTTTCATCTTTTGGAGAGAT 876
QY 238 LeuSerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeu 257

Db 877 CTAGTGAGCATGTAGTTGAGGTTGATGTTCTTCTGACACGTCGTCGACGATGCTC 936
 Qy 258 LeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMet 277
 Db 937 CTGTCTCTTACCATTGCTGAGGTGGAAGAGCGTGGATCCTTACTCTTCCCATCATG 996
 Qy 278 SerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleLysPro 297
 Db 997 AGCAGAAATTCAGAAAAATATAGCAAAAGTCAGAGTTGATTTTATCATCATCAAGCCA 1056
 Qy 298 LeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTyrIleProArgile 317
 Db 1057 TTACCTGATATAGTTGTTCTATGCAGCTTCTCATTTTCCAAAGTATTTGGAACCAAGAATA 1116
 Qy 318 ProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLys 337
 Db 1117 CCATTGGAGCTTGGACATCGTGGTCAGGGAACCTCAACACGACTGCCAAGTAGCTAANA 1176
 Qy 338 ValGlnGluAsnThrIleAlaSerLeuArgAsnAlaSerHisGlyAlaAlaPheVal 357
 Db 1177 GTACAGGAAATATCATCGCTTCTTTAAGAAATGCTGCCAGTCATGGCGCAGCATTTGTA 1236
 Qy 358 GluPheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCys 377
 Db 1237 GAATTTGATGTCACCTTTCAAGGACTTGTGCCCGTGGTGTATCATGACCTCACCTGC 1296
 Qy 378 CysLeuThrMetLysLysPheAspAlaAspProValGluLeuPheGluIleProVal 397
 Db 1297 TGTCTGACCATGAAGAAATATGAGCTGATCCAGTTGATTTGTAATCCAGTA 1356
 Qy 398 LysGluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLys 417
 Db 1357 AAAGAAATTAACATTTGACCAACTCCAGTTATTGAAGCTTTCTCATGTGCTGCAATTA 1416
 Qy 418 SerLysAspArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnPro 437
 Db 1417 ACCAAAGACCGGAAACAATCTTGTATGAGGAGGAAATTTCTTTCTGAAATCAGCCA 1476
 Qy 438 PheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGlu 457
 Db 1477 TTTCTCTCTTAAGATGGTTTTAGAAATCATTTGCCAGAAATGTAGGAATTTAATATAGA 1536
 Qy 458 IleLysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPhe 477
 Db 1537 ATAAATGGATTTGCCAACACAGGGATGGATGGATGGCACTTATCAACATATTTT 1596
 Qy 478 AspMetAsnLeuPheAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArg 497
 Db 1597 GATATGAATGTGTTTTTGGATATATATTTAAAACTGTTTTAGAAAAATTTCTGGGAAGAGG 1656
 Qy 498 ArgIleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsn 517
 Db 1657 AGAATAGTGTCTCTCTTTGATGCAGATATTTGTACATGTTTGGCAGAACAGCAAGC 1716
 Qy 518 LysTyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAsp 537
 Db 1717 AAATATCCCATATATTTTGGCCCAAGGAAAGTCTGATATTTACCCCGAATCATGGAC 1776
 Qy 538 LeuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGly 557
 Db 1777 CTCAGATCTCGGACAAACCCATCTGCAATGTGATTTTGCACAGTTTGAATAATATTTTGGG 1836
 Qy 558 IleLeuValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAla 577
 Db 1837 ATAAATGCCCATCTACGAAAGACCTCTTGAAGAACCCATCTATGTCGAAGAGCAAAAGCT 1896
 Qy 578 LysGlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLys 597
 Db 1897 AAGGGATTGGTCATATTTCTGTGGGGTGTATGATACCAACGATCTGTGAAACAGAGGAA 1956
 Qy 598 LeuLysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGlu 617

Db 1957 CTGAAGGAATTTGGAGTAATAATGGTCTCTAATATATATAGGATATATGATGGATCCCTGAA 2016
 Qy 618 GlnProAsnIlePheGlnValGlnGlnLeuGluArgLeuLysGlnGluLeuProGluLeu 637
 Db 2017 CAGCCAATATATTTCCAAGTTGGACAGTTGGAGCGCTGAAGCAAGAAATTTGCCAGAGCTT 2076
 Qy 638 LysSerCysLeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSer 657
 Db 2077 AAGAACTGTTGTGTCCCACTGTTAGCCACTTTCATCCCTTCTTCTTCTGTGTGGAGCCT 2136
 Qy 658 AspileHisValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
 Db 2137 GATATCACGTGGATGCCAAGCGCATTTGATAGTGTGGAGAAGCT 2181
 RESULT 2
 LOCUS AK050318 3614 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus adult male liver tumor cDNA, RIKEN full-length enriched library, clone: C730037B04 product: hypothetical Glycero-phosphoryl diester phosphodiesterase/Glycosyl hydrolase, starch-binding domain containing protein, full insert sequence.
 ACCESSION AK050318
 VERSION AK050318.1 GI:26093898
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
 REFERENCE 1 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,K., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gscc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissue was provided by William A. Held, Roswell Park Cancer Institute, Department of Molecular and Cellular Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/.

FEATURES

```

source
1..3614
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C730037B04"
/db_xref="taxon:10090"
/clone="C730037B04"
/sex="male"
/cell_type="tumor"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
117..2145
/note="hypothetical Glycero-phosphoryl diester phosphodiesterase/Glycosyl hydrolase, starch-binding domain containing protein (InterPro|IPR004129, InterPro|IPR02044, evidence: InterPro) putative"
polyA_signal
3596..3601
/note="putative"
polyA_site
3614
/note="putative"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3614
Score: 3318.50 Matches: 624
Percent Similarity: 96.89% Conservatives: 31
Best Local Similarity: 92.31% Mismatches: 17
Query Match: 93.82% Indels: 4
DB: 3 Gaps: 1

US-10-047-855-3 (1-672) x AK050318 (1-3614)
QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 117 ATGACACCTTCTCAGGTCACTTTTGAATAAGAGGAACCTCTTTTACAGGAGGTCTTT 176
QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 177 GCAATATGTGGAAGCTGTGATGCCCTGGGAACTCGAAATCCCTCAAAATGCTGTGGCTCT 236
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 237 ATTAATGAAACGACGACAGGACAGTGTGTGTGGAAAGCATGTGCTCTCAATAGA 296

```

```

QY 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 297 GGAGTGTGAGTGAAGTACCGTCTTCTCAGAGGCTGCTTTTGTAGAACCAAGACTATCGGT 356
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisIleuGlnProArgSerIleThr 100
Db 357 GGTCCATGTCAAGTCATAGTTTCAACAGTGGGAGACTCATCTACAAACCAACGATCAATACC 416
QY 101 ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValG 120
Db 417 CCCTTTAGAAAGTGAATTCATTTATGACGATGACGAGTTTGGCATCCACATGGTGTGA 476
QY 120 uThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSe 140
Db 477 AACACTGGATTCCTGGATGGCTTACATGTCAGACTGAATAAGATTCGCTCGCATTTTTC 536
QY 140 rGluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgVal 160
Db 537 TGAGAAACCTCTCTGTTTCAATTAGCAAGAAAGTTCAAAAATCGAGATTTAGGGTAAA 596
QY 160 sLeuThrIleuGluGlyLeuGluGlu-----AspAspAspAspArgValSerProTh 177
Db 597 GCTCACTCGAGGGTCTGGAGGAAGATGAAGATGATGATGACGATAGGTCTCTCCAC 656
QY 177 rValLeuHisLysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPhe 197
Db 657 TGTTCCTTCAAAAATGTCCAACAGCTGGAGATATCTTAAATAGTGACAATGAGTTCAA 716
QY 197 sCysArgHisSerGlnProGluCysGlyTrpGlyLeuGlnProAspArgTrpThrGluTy 217
Db 717 GTGAGGCACCTCACGCCAGAAATGTGGGTATGGCTTTACAGCCCGATCGTTGGACAGAT 776
QY 217 rSerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPhePheGluGluAs 237
Db 777 CAGCATACAGACAATGGAAACCAAGATAATCTGGAGCTCATCTTTGACTTTTGGAGAGA 836
QY 237 pLeuSerGluHisValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLe 257
Db 837 TCTCAGTGAGCATGTAGTTCCAGGTGATGTTCTCTCTCGACACGCTGGGACACGATGCCT 896
QY 257 uLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrProIleMe 277
Db 897 CCTGTCTTCTACCATTCCTGAGAGTGAAGAGCGCTGGAATCCTTACTCTTCCCACAT 956
QY 277 tSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysPr 297
Db 957 GAGCAGAAATTCAGAAATACTATAGCAAAAGTCAGAGTTGATTTTATCATCATCAAGCC 1016
QY 297 oLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArg 317
Db 1017 ATTACCTGGATATAGTTGTTCTATGCACTCTTCATTTTCCAAGTATTGGAAACCAAGAAT 1076
QY 317 eProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAla 337
Db 1077 ACCATTGGAGCTTGGACATCGTGTGAGGAACTCAACAAACGACTGCCAAGTAGCTAA 1136
QY 337 sValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheVa 357
Db 1137 AGTACAGGAAATACATATCGCTTCTTTAAGAAATGTGCCAGTCATGGCGCAGCAATTTGT 1196
QY 357 lGluPheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCy 377
Db 1197 AGAATTTGATGTCCACCTTTCAAAGGACTTTGTGCCGCTGGTGTATCATGACCTCACCTG 1256
QY 377 sCysLeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProVa 397
Db 1257 CTGTCTGACCATGAAGAGGAAATATGAAGCTGATCCAGTTGAAATTTGTTGAAATCCCGAT 1316
QY 397 lLysGluLeuThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLy 417
Db 1317 AAAAGAATTAACATTTGACCAACTCCAGTTATTTGAAAGCTTTCTCATGTGACTGCATTAAA 1376
QY 417 sSerLysAspArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnPr 437

```

```
Db 1377 AACCAAGACCCGGAACAATCTTTGATGAGGAGGAAATTTCTTTCTGAAAAATCAGCC 1436
Qy 437 oPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleG1 457
Db 1437 ATTCTCTCTTAAGATGTTTAAAGATCATCTGCCAGAAATGTAGGATTAAATATAGA 1496
Qy 457 uileLysTrpIleCysGlnArgAspGlyMetTrpAspGlyAsnLeuSerThyTrpPh 477
Db 1497 AATAAAATGGATTGCGCAACAGAGGATGGATGGATGGATGGATGGATGGATGGAT 1556
Qy 477 eAspMetAsnLeuPheLeuAspIleLeuLysThrValLeuGluAsnSerGlyLysArg 497
Db 1557 TGATATGAATGTTTGTGATATATATTTTAAACTGTTTGTAGAAATCTGGGAAGAG 1616
Qy 497 gArgIleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAs 517
Db 1617 GAGATAGTGTCTTCTTTGATGCGAGATATTTGTACATGGTTTGGCAGAGCAGNA 1676
Qy 517 nLysTyrProIleLeuPheLeuThrGlnGlyLysSerGluLysLeuTrpGluLeuMetAs 537
Db 1677 CAAATATCCATATATATTTTTCACCAAGAAAGTGTATATTTTACCCGAACTCATGA 1736
Qy 537 pleuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlu 557
Db 1737 CCTCAGATCTCGCAACACCCATTGCAATGATGTTTGCACAGTTTGAATAATATTTTGG 1796
Qy 557 yIleAsnValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAl 577
Db 1797 GATAAATGCCATATCTAGAGACTCTCTTAGAACCCATCTATGTCCAGAGGCAAAAGC 1856
Qy 577 aLysGlyLeuValIlePheCysTrpGlyAspThrAsnAspProGluAsnArgArgLys 597
Db 1857 TAAGGGATGGTGCATATCTGCTGGGTGATGATACCAACGATCTCTGAAACAGAAAGAA 1916
Qy 597 sleuLysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGl 617
Db 1917 ACTGAAGGAATTTGAGTAATATGGTCTAATATATGATAGGATATATGATGGATCCCTGA 1976
Qy 617 uGlnProAsnIlePheGlnValGlnLeuGluArgLeuLysGlnGluLeuProGluLe 637
Db 1977 ACAGCCAAATATATTCAGTGAGCAGATTGAGGCGCTGAGCAAGAAATGCCAGACT 2036
Qy 637 uLysSerCysLeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSe 657
Db 2037 TAAGAACTGTTTGTCTCCACTGTTAGCCACTTCATCCCTCTCTTCTTGTGTGGAGCC 2096
Qy 657 rAspIleHisValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2097 TGATATCCAGCTGGATGCCAAGCGATTGATAGTGTGGAGAACGCT 2142

RESULT 3
BC006887 3248 bp mRNA linear HTC 19-NOV-2003
LOCUS BC006887
DEFINITION Mus musculus RIKEN cDNA 2310032D16 gene, mRNA (cDNA clone
IMAGE:3596618), containing frame-shift errors.
ACCESSION BC006887
VERSION BC006887.1 GI:14711816
KEYWORDS HTC.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3248)
AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
```

```
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3248)
Strausberg, R.
Direct Submission
Submitted (27-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgobcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisege, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 16 Row: 1 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: frame shifted.
FEATURES
source
Location/Qualifiers
1..3248
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3596618"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal Carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DHI0B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3248
Score: 3150.00 Matches: 598
Percent Similarity: 93.19% Conservative: 31
Best Local Similarity: 88.59% Mismatches: 18
Query Match: 89.06% Indels: 29
DB: 3 Gaps: 2
US-10-047-855-3 (1-672) x BC006887 (1-3248)
Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 127 ATGACACCTCTTCAGTTCACCTTTTGAATAAGAGAACTCTTTTACCAAGAGAGTCTTT 186
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
```

Db	187	GCAATATGTGGAGCTGTGATGCCCTGGGAAACTGGAATCCTCAAAATGCTGTGGCTCTT	246
Qy	41	LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg	60
Db	247	ATTAATGAANAACGACACAGACAGACAGTGTGTGTGGAAAGCAGTATTGCTCTCAATAGA	306
Qy	61	GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly	80
Db	307	GGAGTGTGAGTGAAGTACCGCTACTTCAGAGGCTGCTTTTAGAA-----	351
Qy	81	GlyProCysGlnValIleValHisLysTyrGluThrHisLeuGlnProArgSerIleThr	100
Db	351	-----	351
Qy	101	ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu	120
Db	352	CCAAA-GAAAGTGAATCATATTGACGATGGACAGTTTGGCATCCACAATGGTGTGA	410
Qy	121	ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer	140
Db	411	ACACTGGATTCTGGATGGCTTACATCTCAGACTGAAATTAAGATTTGCGTCTGCATTTTTCT	470
Qy	141	GluLysProProValSerIleThrLysLysLeuLysLysSerArgPheArgValLys	160
Db	471	GAGAAACCTCTGTTCATTTAGCAAGAAAGATTCAAAAAATCCAGATTTAAGGTTAAG	530
Qy	161	LeuThrLeuGluGlyLeuGluGlu-----AspAspAspArgValSerProThr	177
Db	531	CTCACACTCGAGGCTGTGGAGGAAGTAGAGATGATGACGATAAGGTCTCTCCACT	590
Qy	178	ValLeuHisLysMetSerAsnSerLeuGluIleSerLysLysSerAspAsnGluPheLys	197
Db	591	GTTCTTCACAAAATGTCCAAACAGCTGGAGATATCCTTAATAGTGACAAATGAGTTCAG	650
Qy	198	CysArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyr	217
Db	651	TGCAGGCACCTCACAGCCAGAAATGTGGGTATGGCTTACAGCCCATCGTTGGACAGATAC	710
Qy	218	SerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAsp	237
Db	711	AGCATACAGACAAATGGAAACCAGATAATCTGGAGCTCATCTTTGAGGAAGAT	770
Qy	238	LeuSerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeu	257
Db	771	CTCAGTGAGCATGTAGTTTCAGGTGTGTCTTCTCGACACGTGGGCACAGATGCCCTC	830
Qy	258	LeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMet	277
Db	831	CTGTCTCTACCATGTCTGAGTGGAGAACGCCCTGGAAATCCTTACTCTTCCCATCATG	890
Qy	278	SerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleLysPro	297
Db	891	AGCAGAAATTCAGNAAATCATAGGCCAAAGTCAGAGTTGATTTTATCATCATCAAGCCA	950
Qy	298	LeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIle	317
Db	951	TTTACCTGGATATAGTTGTTCTATGCAAGCTCTTCAITTTTCCAAAGTATTGGAAACCAAGAATA	1010
Qy	318	ProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrThrAlaGlnLeuAlaLys	337
Db	1011	CCATTTGGACGTTGGACATCGTGGTGGAGGAACCTCAACAGACTGCCAGCTAGCTTAA	1070
Qy	338	ValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheVal	357
Db	1071	GTACAGGAAATATCATTCGTTCTTTAGAAATGCTGCCAGTCATGCGCAGCATTTGTA	1130
Qy	358	GluPheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCys	377
Db	1131	GAAATTGATGTCCACTTTCAAAGACATTTGTGCCCGTGGTGTATCATGACCTCACTGC	1190
Qy	378	CysLeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProVal	397
Db	1191	TGTCGTACCATGAAGAGGAAATATGAGAGCTGATCAGTTGAATGTTTGAATCCAGTA	1250

398	QY	lysGluLeuThrPheAspGlnLeuGlnLeuLeuLeuLeuHisValThrAlaLeuLys	417
1251	Db	AAAGAAATTAACATTTGACCACTCCAGTATTAGAGCTTTCTCATGTGACATGCAATAAAA	1310
418	QY	SerLysAspArgLysGluSerValValGlnGluGluAsnSerPheSerGluAsnGlnPro	437
1311	Db	ACCAAAGACCGGAACAATCTTTGTATGAGGAGGAAAATTTCTTTCTTGAAAATCAGCCA	1370
438	QY	PheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGlu	457
1371	Db	TTTTCTCTCTTAAGATGGTTTTAGAAATCATTTGCCAGAAAATGTAGGATTTAAATATAGAA	1430
458	QY	IleLysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTrpPhe	477
1431	Db	ATAAAATGGATTTGCCAACACAGGGATGGAGTATGGGATGGCAACTATCAACATATTTT	1490
478	QY	AspMetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArg	497
1491	Db	GATATGAATGTGTTTTGGATATAATTTTAAAAA CTGTTTTAGAAAAATTC TGGGAAGAGG	1550
498	QY	ArgIleValPheSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsn	517
1551	Db	AGBATAGTCTTTCTTTTCATGCAGATAATTGTACAATGGTTCCGGCAGAAAGCAGAAC	1610
518	QY	LysTyProIleLeuPheLeuThrGlnGlyLysSerGluIleTyProGluLeuMetAsp	537
1611	Db	AAATATCCCATATATATTTTGAACCAAGAAAGTCTGATATTATTTACCCCAACTCATGGAC	1670
538	QY	LeuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGly	557
1671	Db	CTCAGATCTCGGACAACACCATTCGCAATGAGTTTTGCACAGTTTGAAATAATTTTGGGG	1730
558	QY	IleAsnValHisThrGluAspLeuLeuArgAsnProSerTyIleGlnGluAlaLysAla	577
1731	Db	ATAAATGCCATCTACTGAAGACCTCTCTAGAAACCCATCCTATGTCCAAAGGGCAAAAGCT	1790
578	QY	LysGlyLeuValIlePheCysTyTrpGlyAspAspThrAsnAspProGluAsnArgArgLys	597
1791	Db	AAGGGATGTGTCATATTCCTGCTGGGGTGATGATACCACACGATCCTGAAAACAGAGGAAA	1850
598	QY	LeuLysGluLeuGlyValAsnGlyLeuIleTyArgPheValProSerLeuCysGlyGluSer	617
1851	Db	CTGAGGAAATTTGGAGTAAATGCTCTAAATATATGATAGGATATATGATTGGATGCCCTGAA	1910
618	QY	GlnProAsnIlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeu	637
1911	Db	CAGCCAATATATTTCCAAGTGGAGCAGTGTGGAGCGCTGAAAGCAGAAATTTGCCAGAGCTT	1970
638	QY	LysSerCysLeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSer	657
1971	Db	AAGAAATGTTTTGTGTCCACTGTGTAGCCACTTCAATCCCTTCTCTTCTTCTGTGTGGAGCCT	2030
658	QY	AspIleHisValAspAlaAsnGlyIleAspAsnValGluAsnAla	672

2

TITLE
JOURNAL

High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED
10349636

REFERENCE
AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED
11042159

REFERENCE
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaehiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861

REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

TITLE
JOURNAL

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3578)

REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/

FEATURES
source

Location/Qualifiers
1. 3578
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6030473012"
/db_xref="taxon:10090"
/clone="6030473012"

/sex="male"
/tissue type="testis"
/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="13 days embryo"
205_-2109
/note="hypothetical Glycerophosphoryl diester
phosphodiesterase/Glycosyl hydrolase, starch-binding
domain containing protein (InterPro|IPR004129,
InterPro|IPR02044, evidence: InterPro)
putative"
3560..3565
/note="putative"
3578
/note="putative"

misc_feature

polyA_signal

polyA_site

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3578
Score: 3066.00 Matches: 585
Percent Similarity: 90.83% Conservative: 29
Best Local Similarity: 86.54% Mismatches: 17
Query Match: 86.68% Indels: 46
DB: 3 Gaps: 2

US-10-047-855-3 (1-672) x AK031658 (1-3578)

Qy 1 MetThrProSerGlnValAlaPheGluileArgGlyThrLeuLeuProGlyGluValPhe 20
Db 205 ATGACACTTCTCAGGTCACCTTTTGAATAAGAGAGAACTCTTTTACCAGGAGAGTCTTT 264
Qy 21 AlaileCysGlySerCysAspAlaLeuGlyAsn-TTpAsnProGlnAsnAlaValAlaLe 40
Db 265 GCAATATGTGGAAGCTGTGATGCTCCCTGGAAACCTCGAATCCTCAAAATGCTGTGGCTCT 324
Qy 40 uLeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrileValLeuSerAr 60
Db 325 TATTAAATGAAACACGACGAGACAGTGTGTGTGGAAAGCAGTGTGCTCTCAATAG 384
Qy 60 qGlyValSerValGlnTrpArgTyrPheLysGlyTyrPheLeuGluProLysThrileGl 80
Db 385 AGGAGTGTGAGTGAAGTACCGCTACTTTCAGAGGCTGCTTTTATAGACCAAGACTATCGG 444
Qy 80 YGlyProCysGlnValleValHisLysTrpGluThrHisLeuGlnProArgSerileTh 100
Db 445 TGGTCCATGTCAAGTCATAGTTCAAGTGGGAGACTCATCTACCAACCACGATCAATAAC 504
Qy 100 rProLeuGluSerGluilellelleAspAspGlyGlnPheGlylleHisGlyValGl 120
Db 505 CCCTTTAGAAAGTGAATTCATTATTGACGATGGACAGTGTGGCATCCAC----- 553
Qy 120 uThrLeuAspSerGlyTrpLeuThrCysGlnThrGluileArgLeuArgLeuHisTyrSe 140
Db 553 ----- 553
Qy 140 rGluLysProProValSerileThrLysLysLysLeuLysLysSerArgPheArgVally 160
Db 554 -----AGGTA-AA 560
Qy 160 sLeuThrleuGluGlyLeuGluGlu-----AspAspAspArgValSerProth 177
Db 561 GCTCACACTCGAGGCTCTGGAGGAAGATGAAGATGATGATGATGATGATGATGATGATGAT 620
Qy 177 rValLeuHisLysMetSerAsnSerLeuGluileSerLeuileSerAspAsnGluPheLy 197
Db 621 TGTTCCTTCAAAAATGTCACACGCTGGAGATATCTTAAAGTGCATGATGATGATGATGAT 680
Qy 197 sCysArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTy 217
Db 681 GTGCAGGCACTCACAGCCAGAAATGTGGGTATGGCTTACAGCCCGATCGTTGGACAGAGTA 740
Qy 217 rSerileGlnThrMetGluProAspAsnLeuGluilellePheAspPheGluGluAs 237
Db 741 CAGCATACAGCAATGGAACCCAGATAATCTGGAGCTCATCTTTGACTTTTGTGGAAGA 800

QY	237	pLeuSerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysIle	257	QY	597	sLeuLysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProG1	617
Db	801	TCCTCAGTGAGCATGTAGTTCAGGCGTATGTTCTTCTCGGACACGTCGACGATGCCT	860	Db	1881	ACTGAGGAAATTTGGAGTAAATGCTCTAATATATATAGTAGGATATATGATTTGGATGCCTGA	1940
QY	257	uLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMe	277	QY	617	uGlnProAsnIlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLe	637
Db	861	CCTGCTCTTACCATTGCTGAGAGTGGAAGCGCTGGAATCCTTACTCTTCCCATCAT	920	Db	1941	ACAGCCAAATATATATCCAAAGTGGAGCAGTTGGAGCGCTGAAGCAAGAAATTCGACAGCT	2000
QY	277	tSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysPr	297	QY	637	uLysSerCysLeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSe	657
Db	921	GAGCAGAAATTCAGAAATACTATAGCAAGTCAGAGTTGATTTTATCATCATCATCAAGCC	980	Db	2001	TAAGAACTGTTTGTGTCCACTGTTAGCCACTTCATCCCTCTTCTTCTGTGTGGAGCC	2060
QY	297	oLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIl	317	QY	657	rAspIleHisValAspAlaAsnGlyIleAspAsnValGluAsnAla	672
Db	981	ATTACCTGGATATATGTTCTATGCAGTCTTCAATTTTCCAAAGTATTTGGAACCAAGAAAT	1040	Db	2061	TGATATCCACGTCGATGCGCAACGCGATGATAGTGTGGAGAACGCT	2106
QY	317	eProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLys	337	RESULT 5			
Db	1041	ACCATTGGACGTTGCATCCTTCAAGGACTTTGTGCGCGTGTATCATGACCTCACCTG	1100	LOCUS	AK009563	2604 bp mRNA linear HTC 03-APR-2004	
QY	337	sValGlnGluAsnThrIleAlaSerLeuAtqAsnAlaAlaSerHisGlyAlaAlaPheVa	357	DEFINITION		Mus musculus adult male tongue cDNA, RIKEN full-length enriched	
Db	1101	AGTACAGGAAATACTATCGCTTCTTTAAGAAATGCTGCCAGTCATGCGCAGCAATTTGT	1160			library, clone:2310032D16 product:hypothetical Glycero-phosphoryl	
QY	357	lGluPheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCy	377			diester phosphodiesterase/Glycosyl hydrolase, starch-binding domain	
Db	1161	AGNATTGATGTCACCTTTCAGGACTTTGTGCGCGTGTATCATGACCTCACCTG	1220	ACCESSION	AK009563	containing protein, full insert sequence.	
QY	377	sCysLeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProVa	397	VERSION	AK009563.1		
Db	1221	CTGTCTGACCATGAAGAGAAATATGAAGCTGATCCAGTTCGAAATGTTGAAATCCCGAGT	1280	KEYWORDS	GI:12844432		
QY	397	lLysGluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLys	417	SOURCE	HTC: CAP trapper.		
Db	1281	AAAGAAATTAACATTGACCACTCCAGTTATGAGCTTCTCATGTGACTGCATTAA	1340	ORGANISM	Mus musculus (house mouse)		
QY	417	sSerLysAspArgLysGlnValValGlnGluAsnSerPheSerGluAsnGlnPr	437	REFERENCE	Mus musculus		
Db	1341	RACCAAGACCGAACAACATCTTTGATGAGGAGGAAATTTCTTTCTGAAATTCAGCC	1400	AUTHORS	Mus musculus		
QY	437	oPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGl	457	TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Db	1401	ATTTCCTCTTAAGATGGTTTGAATCAATTCGACGAAATATGAGATTTAATATAGA	1460	JOURNAL	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
QY	457	uIleLysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTrpPh	477	MEDLINE	Carninci, P. and Hayashizaki, Y.		
Db	1461	ATAATAATGGATTTGCCAACACAGGATGGATATGGATGGCACTTATCAACATATTT	1520	PUBMED	High-efficiency full-length cDNA cloning		
QY	477	eAspMetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysAr	497	REFERENCE	Meth. Enzymol. 303, 19-44 (1999)		
Db	1521	TGATATGAATGTTTTCGATATAATTTTAAAAACTGTTTGAAGAAATTCCTGGAGAG	1580	AUTHORS	10349636		
QY	497	gArgIleValPheSerSerPheAspAlaAspIleCysThrMetValArgLysGlnAs	517	TITLE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
Db	1581	GAGATAGTGTTCCTCTTCCTTTGATGCAGATATTTGTACAAATGTTTCGACAGAGCAAA	1640	JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to		
QY	517	nLysTyrProIleLeuPheLeuThrGlnLysSerGluIleTyrProGluLeuMetAs	537	MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes		
Db	1641	CAAAATCCCATATATTTTGTACCAAGGAAGTCTGATATTTACCCGCACTCATGGA	1700	PUBMED	Genome Res. 10 (10), 1617-1630 (2000)		
QY	537	pLeuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlu	557	REFERENCE	20499374		
Db	1701	CCTCAGATCTCGACACACCCCATTCGATGATTTTGCACAGTTTGAAGAAATATTTGGG	1760	AUTHORS	11042159		
QY	557	yIleAsnValHisThrGluAspLeuArgAsnProSerTyrIleGlnGluAlaLysAl	577	TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
Db	1761	GATAAATGCCCACTACTGAAGACCTCTCTAGAACCCCATCCTATGTCGAGAGGCAAGC	1820	JOURNAL	Konno, H., Akiyama, J., Nishi, K., Kusunai, T., Tashiro, H., Itoh, M.,		
QY	577	aLysGlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgLys	597	MEDLINE	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Mishine, T., Harada, A.,		
Db	1821	TAAGGGATTCGTATATTTCTGCTGGGTGATGATACCAAGCTCTCTGAAAACAGAGGAA	1880	PUBMED	Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		

Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayaishizaki, Y.

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

PLEASE VISIT OUR WEB SITE (<http://genome.gsc.riken.jp/>) FOR FURTHER DETAILS.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence (5'-GAGAGAGAGATTCGAGTTAATTAATTAATCCCTCCCCC 3'). cDNA was cleaved with XhoI and SmaI. Cloning sites, 5' end: XhoI; 3' end: SmaI. Host: SOLR.

FEATURES

source

```
1. 2604
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:2310032D16"
/db_xref="taxon:10090"
/clone="2310032D16"
/sex="male"
/tissue_type="longue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
94. 1449
```

CDS

```
/note="unnamed protein product; hypothetical
glycerophosphoryl diester phosphodiesterase/glycosyl
hydrolase, starch-binding domain containing protein
(InterPro|IPR004129, InterPro|IPR002044, evidence:
InterPro)
putative"
/codon_start=1
/protein_id="BAB26361.2"
/db_xref="GI:26365211"
```

```
/translation="MEPNLELIPDFPELSEHVQDVLPGHVGTCALSLSTIAES
GRSAGILTPIMSRNRTIGKRVDFIIIKPLPGYSCMSQSSFSKYWKPRIPDLVGH
RGANSTTAKLAQNTIASLRNASHGAFFVEFVHLSKDFVPPVYHDLTCLDM
KRKYEADPVELFEI PVKELTDLQOLLKSHVLTALKDKRKQSLYEENFEENQFPF
SLKWLSESLPNVGNIEIKWTICQHRDGVNDNLSTYFDMNVFLDIILKTVLENSGKR
RIVSFSEADICTWVQKQKYPILFLTQKSDIYPELMDLRSRTTPAMSAQFENI
SGINAHTELLNPSYQEKAKGLVIFCWGDDNDPENRRKLEEFVNGVGLIYDIYD
WMPEQPNIFQVEQLERLKOELPELKNCLICTVSHPIFSSFCVPEPDHVDANGIDSVEN
A"
```

ORIGIN

Alignment Scores:	6.16e-262	Length:	2604
Pred. No.:	2400.00	Matches:	448
Score:	97.51%	Conservative:	22
Best Local Similarity:	92.95%	Mismatches:	12
Query Match:	67.85%	Indels:	0
DB:	3	Gaps:	0

US-10-047-855-3 (1-672) x AK009563 (1-2604)

QY	191	IleSerAspAsnGluPheLysCysArgHisSerGlnProGluCysGlyTyrGlyLeuGln	210
DB	1	ATAAGTGACATAGATTCAAGTCAGGCACCTCACAGCCAGAAATGGGTATGGCTTACAG	60
QY	211	ProAspArgTTPThrGluTyrSerIleGlnThrMetGluProAspAsnLeuGluLeuile	230
DB	61	CCCGATCGTTGGACAGAGTACAGCATACAGCAATGGACACAGATATCTGGAGCTCATC	120
QY	231	PheAspPheGluGluAspLeuSerGluHisValValGlnGlyAspAlaLeuProGly	250
DB	121	TTTGACATTTTGGAGGAAGATCTCAGTGAGCATGTAGTTTCAGGGTGTATGTTCTTCTCGGA	180
QY	251	HisValGlyThrAlaCysLeuLeuSerSerThrIleAlaGluSerGlyLysSerAlaGly	270
DB	181	CACGTGGGCACAGCATGCTCTCTCTTCTACCATTCGTGAGAGTGGGAAGCGCTCGA	240
QY	271	IleLeuThrLeuProIleMetSerArgAsnSerArgLysThrIleGlyLysValArgVal	290
DB	241	ATCCTTACTCTTCCCATCATGAGCAGAAATTCAGAAAATACTATAGCCAAAGTCAGATT	300
QY	291	AspTyrIleIleIleLysProLeuProGlyTyrSerCysAspMetLysSerSerPheSer	310
DB	301	GATTTATCATCATCAAGCCATTACCTGGATATAGTTGTTCTATGCAGCTCTTCATTTCC	360
QY	311	LysTyrTrpLysProArgIleProLeuAspValGlyHisArgGlyAlaGlyAsnSerThr	330
DB	361	AAGTATTGGAAACCAAGAAATACCATTTGGACGCTGGACATCGTGGCAGGGAACCAACA	420
QY	331	ThrThrAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAla	350
DB	421	ACGACTGCCAAGTAGCTAAAGTACAGAAATACATTCGCTTCTTTTAAAGAAATGCTGCC	480
QY	351	SerHisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheValProVal	370
DB	481	AGTCATGCGCAGCATTTGTAGATTTGATGTCCACCTTTCAGAGGACTTTGTGCCCCGTG	540
QY	371	ValTyrHisAspLeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspProVal	390
DB	541	GTGTATCATGACCTCACCTGCTGTGACCATGAAGAGAAATATGAAGCTGATCCAGTT	600
QY	391	GluLeuPheGluIleProValLysGluLeuThrPheAspGlnLeuGlnLeuLysLeu	410
DB	601	GAATGTTTGAATCCCAAGTAAAGAAATTAACATTTGACCAACTCCAGTTATTAAGCTT	660
QY	411	ThrHisValThrAlaLeuLysSerLysAspArgLysGluSerValValGlnGluAsn	430
DB	661	TCTCATGTGATGTCATTAATAAACCAAGACCGGAAACAATCTTTGTATGAGGAGGAAAT	720
QY	431	SerPheSerGluAsnGlnProPheProSerLeuLysMetValLeuGluSerLeuProGlu	450
DB	721	TTCTTTTCTGAAATCAGCCATTTCTTCTTAAGATGGTTTGTAGAAATCATTTGCCAGAA	780
QY	451	AspValGlyPheAsnIleGluIleLysTrpIleCysGlnGlnArgAspGlyMetTrpAsp	470
DB	781	AATGTAGGATTAATAATAGAAATAAATGGATTTGCCAACACAGGATGGAGATGGGAT	840
QY	471	GlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeuAspIleLeuLysThrVal	490
DB	841	GGCACTTATACACATATTTGATATGATGTGTTTGGATATATATTTTAAAAACTGTT	900
QY	491	LeuGluAsnSerGlyLysArgArgIleValPheSerSerPheAspAlaAspIleCysThr	510
DB	901	TTAGAAAATTTCTGGGAAGAGGAGAAATAGTGTGTTTCTTTTGTATGACAGATATTGTACA	960
QY	511	MetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeuThrGlnGlySerGlu	530
DB	961	ATGTTGCGCAGAGACAGACAAATATATTTTGGACCAAGGAAGTCTGAT	1020
QY	531	IleTyrProGluLeuMetAspLeuArgSerArgThrThrProIleAlaMetSerPheAla	550
DB	1021	ATTTACCCCGAAGTCTAGGACCTCAGATCTCGGACACACCCATTCGCAATGAGTTTGA	1080
QY	551	GlnPheGluAsnLeuLeuGlyIleAsnValHisThrGluAspLeuLeuArgAsnProSer	570


```

US-10-047-855-3 (1-672) x AK049491 (1-1618)
QY 235 GluGluAspLeuSerGluHisValValGlnGlnGlyAspAlaLeuProGlyHisValGlyThr 254
DB 3 GAGGAAGATCTCAGTGAGCATGTAGTTCAGGGTGATGTTCTTCCTGGACACGTGGGCACA 62
QY 255 AlaCysLeuLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeu 274
DB 63 GCATGCTCTCTCTTACCATTCGTGAGAGTGAAGAGCGCTCGAATCCTTACTCTT 122
QY 275 ProIleMetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIle 294
DB 123 CCATCATGAGCAGAAATTCAGAGAAACTATAGCAAGTCCAGAGTTCATTTCCAGTATTCATC 182
QY 295 IleLysProLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLys 314
DB 183 ATCAAGCCATTAACCTGGATATAGTTCATGCGAGTCTTCATTTCCAGTATTCAGTAA 242
QY 315 ProArgIleProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGln 334
DB 243 CCAAGAAATACCATTCGACATCGTGTGCGAGGAACCTCAACAACGACTGCCAAG 302
QY 335 LeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaSerHisGlyAla 354
DB 303 CTAGCTAAAGTACAGGAATACTATCGCTCTTTTAAGAAATGCTGCCAGTCATGCCCA 362
QY 355 AlaPheValGluPheAspValHisLeuSerLysAspPheValProValValTyrHisAsp 374
DB 363 GCATTTGTAGATTTGATGTCCACCTTTCAAAGGACTTGTGCGCGTGGTGTATCATGAC 422
QY 375 LeuThrCysCysLeuThrMetLysLysPheAspAlaAspProValGluLeuPheGlu 394
DB 423 CTCACCTGCTGCTGACCATGAAGAGAAATATGAAGCTGATCCAGTTGAATGTTTGAA 482
QY 395 IleProValLysGluLeuThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThr 414
DB 483 ATCCAGTAAAGAAATTAACATTTCACCAACTCCAGTTTATTGAAGCTTCTCATGTGACT 542
QY 415 AlaLeuLysSerLysAspArgLysGluSerValValGlnGluAsnSerPheSerGlu 434
DB 543 GCATTAATAAACCAAGACCGGAACAATCTTTGTATGAGGAGAAATTTCTTTCTGAA 602
QY 435 AsnGlnProPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPhe 454
DB 603 AATCAGCCATTTCTCTCTTAAGATGGTTTATGAATCATTCGCCAGAAATCTAGCATTT 662
QY 455 AsnIleGluIleLysTrpIleCysGlnArgAspGlyMetTrpAspGlyAsnLeuSer 474
DB 663 AATATACAAATAAATGGAATTTGCCAACACAGCGGATGGAGTATGGGATGGCAACTATCA 722
QY 475 ThrTyrPheAspMetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSer 494
DB 723 ACATATTTTCGATATGAATGTGTTTTTGGATATAATTTTAAACACTGTTTCACAAATTTCT 782
QY 495 GlyLysArgArgIleValPheSerPheAspAlaAspIleCysThrMetValArgGln 514
DB 783 CGGAACAGGAGAAATAGTGTCTTCTTTTGTATGATGATATTTGTACATGTTCTCGGACG 842
QY 515 LysGlnAsnLysTyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGlu 534
DB 843 AAGCAGAAACAATATCCCATATATTTTTCAGCAAGGAAAGTCTCATATTTTACCCCGAA 902
QY 535 LeuMetAspLeuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsn 554
DB 903 CTCATGGACCTCAGATCTGGCAACACCCATTCGAATGAGTTTTCACAGTTTGAATAAT 962
QY 555 LeuLeuGlyIleAsnValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGlu 574
DB 963 ATTTTGGGATTAATGCCCATACTAGACCTCTCTTAGAACCCATCTCATGTCCAGAG 1022
QY 575 AlaLysAlaLysGlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAsn 594

```

```

DB 1023 GCAAAAGCTAAGGATTTGGTTCATATTTCTGCTGGGGTGATGATACCAACGATCCTCGAAAC 1082
QY 595 ArgArgLysLeuLysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrp 614
DB 1083 AGAAGGAACTGAAGGATTTGGAGTAATGCTCTAATATATATAGGATATATGATTGG 1142
QY 615 MetProGluGlnProAsnIlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeu 634
DB 1143 ATGCTGTGAACGCCAAATATATTTCCAGTGGAGCAGTTGGAGCGCCTGAAGCAAGATTG 1202
QY 635 ProGluLeuLysSerCysLeuCysProThrValSerArgPheValProSerSerLeuCys 654
DB 1203 CGAGAGCTTAAGAACTGTTGTGCTCCACTGTAGCCACTTCATCCCTCTCTTCTTCTGT 1262
QY 655 GlyGluSerAspIleHisValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
DB 1263 GTGGAGCCTGATATCCAGTGGATGCCAAGCGCATTTGATGTGGAGAACGCT 1316

RESULT 7
AK051728 2228 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
DEFINITION enriched library, clone:D130067011 product:hypothetical
Glycerophosphoryl diester phosphodiesterase/Glycosyl hydrolase,
starch-binding domain containing protein, full insert sequence.
ACCESSION AK051728.1 GI:26342153
VERSION AK051728.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE PUBMED 10349636
REFERENCE
2
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE PUBMED 11042159
REFERENCE
3
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE PUBMED 11076861
REFERENCE
4
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2228)
JOURNAL PUBMED
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

```

Hayaashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayaashizaki, Y.

TITLE
JOURNAL
Submitted (16-JUL-2001) Yoshihide Hayaashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
Location/Qualifiers
1..2228
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D130067011"
/db_xref="taxon:10090"
/clone="D130067011"
/tissue_type="spinal ganglion"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="12 days embryo"
103..1293
/note="unnamed protein product; hypothetical
Glycophosphoryl diester phosphodiesterase/glycosyl hydrolase, starch-binding domain containing protein (interPro|IPR0041129, InterPro|IPR020044, evidence: putative)
/codon_start=1
/protein_id="BAC34739.1"
/db_xref="GI:26342154"
/translation="MSRSRKTIGKRVDFIILKPLGYSQSMOSFSKYWKPRIPLD VGRGAGNSTTTAKLAKVOENTIASLRNASHGFAVFDFVHLSKDFVVPVYHDLTCC LTMKRYADPVELEIIPVKLTDPQLKLSHVTLTKDKRQSLYEENFFSENG PPPLKVLSELPENGVNIEIKWICQHRDGVWGNLSYFDNMVFLDILKTVLENS GKRRIVSFSDADICTMVROKONKYPILFLTGQKSDIYPELMDLRSLRTPIAMGFAOF ENILGINARTEDLLENPSVQAKAGLVIICWGDNDPENRRKLKEFGVNGLIYDR IYDWMPEQNIIFQVQLERLKLQELFELKNCILPTVSHFIPSSFVCEPDIHVDANGIDS VENA"

ORIGIN

Alignment Scores:
Pred. No.: 5,98e-230 Length: 2228
Score: 2119.00 Matches: 398
Percent Similarity: 97.44% Conservative: 21
Best Local Similarity: 92.56% Mismatches: 11
Query Match: 59.91% Indels: 0
DB: 3 Gaps: 0

US-10-047-855-3 (1-672) x AK051728 (1-2228)

Qy 243 ValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIle 262
Db 1 GTTCAGGGTGATGTTCTTCCTGGACACGTGGGCACAGCATGCGCTCTCTTCTACCA 60
Qy 263 AlaGluSerGlyIysSerAlaGlyIleLeuThrIleuProIleMetSerArgAsnSerArg 282
|||||:::|||||

Db 61 GCTGAGAGTGGAGAGCGCTGGTAATCTTACTCTTCCCATCATGACGACAGAAATTCGAGA 120
Qy 283 LysThrIleGlyIysValAlaGlyValAspTyrIleIleIleIleIleIleIleIleIleIle 302
121 AAAACTATAGCAAGAGTCAGAGTTGATTTATCATCATCAAGCCATTTACCTGGATATAGT 180
Db 303 CysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAspValGly 322
181 TGTTCATGTCAGTCTTCATTTTCCAAAGTATTGGAAACCAAGAAATACCATTTGGACGTTGGA 240
Db 323 HisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaIysValGlnGluAsnThr 342
241 CATCTGTTGTCAGGGAATCAACACGACTCCCAAGCTAGCTAAAGTACACAGAAATAACT 300
Qy 343 IleAlaSerLeuAlaGlnAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAspValHis 362
301 ATCGCTCTCTTAAGAAATGCTGCCAGTCATGCGCAGCATTTGTAGATTTGATGTCCAC 360
Db 363 LeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMetLys 382
361 CTTTCAAGGACTTTGTGCGCGTGTGTATCATGACCTCACCTGTCTGTACCATGAAG 420
Qy 383 LysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThrPhe 402
421 AGGAAATATGAAGCTGATCCAGTTGAAATGTTTGAATCCCAAGTAAAGAAATTAACATTT 480
Db 403 AspGlnLeuGlnLeuLeuLeuLeuLeuHisValThrAlaLeuLysSerLysAspArgLys 422
481 GACCAACTCCAGTTATTGAAGCTTTCTCATGTGACTGCATTAATAAACCAAGACCGGAAA 540
Qy 423 GluSerValValGlnGlnGluAsnSerPheSerGluAsnGlnProPheProSerLeuLys 442
541 CAACTCTTGTATGAGGAGGAAATTTCTTTCTGAAATCAGCCATTTCTCTCTCTTAAG 600
Db 443 MetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleIleIleIleIle 462
601 ATGCTTTTGAATCATTTGCGAGAAATGTAGGATTTAATATAGAATAAAATAGGATTTGC 660
Qy 463 GlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPhe 482
661 CAACACAGGATGGATGAGTATGGATGGCACTTATCAACATATTTTATGATGAATGTGTTT 720
Db 483 LeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSer 502
721 TTGGATATAATTTTAAAACTGTTTTAGAAAATTTCTGGGAAGAGGAGAGATAGTGTTCCT 780
Qy 503 SerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeu 522
781 TCTTTTGTATGCAGATATTTGTACAATGGTTCGCGAGAGAGCAACAAATATCCCATATTA 840
Db 523 PheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThr 542
841 TTTTGGACCAAGAAAGTCTGATATTTACCCGAACTCATGGACCTCAGATCTCGGACA 900
Qy 543 ThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThr 562
901 ACACCAATTCGAATGAGTTTTGACAGATTTTGAATAATTTTGGGGATAAATGCCATACT 960
Db 563 GluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaIysAlaLysGlyLeuValIle 582
961 GAAGACCTCTCTAGAAACCCATCTTATGTCACAGAGCAAAAGCTAAGGATTTGGTCATA 1020
Qy 583 PheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGly 602
1021 TTCTCTGGGGTGATGATACCAAGATCTCTGAAAAACAGAGAAACTGAAGGAATTTGGA 1080
Db 603 ValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePhe 622
1081 GTAAATGGTCTAATATATATAGTAGATATATGATTTGGATGCTGTAACACGCAAAATATATTC 1140
Qy 623 GlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCysLeuCys 642
1141 CAAGTGGAGCAGTTGGAGCGCTGAAGCAAGAAATTGCACAGCTTAAAGAACTGTTGTGT 1200

QY 643 ProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHisValAsp 662
 |||||
 Db 1201 CCACCTGTTAGCCACTTCATCCCTTCTTCTTGTGTGGAGCCTGATATCCACGTGGAT 1260

QY 663 AlaAsnGlyIleAspAsnValGluAsnAla 672
 |||||
 Db 1261 GCCAACGGCAATTGATAGTGTGGAGAACGCT 1290

RESULT 8
 BX344123 1006 bp mRNA linear EST 07-APR-2004
 LOCUS BX344123 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1059YA01 5-PRIME, mRNA sequence.

ACCESSION BX344123
 VERSION BX344123
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1006)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 1, 2003 this sequence version replaced gi:30311247.

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 9838.r

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?8=CS1A10152E01Q1&c=9838.r>.

FEATURES
 source
 1..1006
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1059YA01"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Alignment Scores:
 Pred. No.: 8,33e-168 Length: 1006
 Score: 1570.00 Matches: 311
 Percent Similarity: 97.50% Conservative: 1
 Best Local Similarity: 97.19% Mismatches: 8
 Query Match: 44.39% Indels: 3
 DB: 5 Gaps: 0

US-10-047-855-3 (1-672) x BX344123 (1-1006)

QY 344 AlaSerLeuArgAsnAlaAalaSerHisGlyAlaAlaPheValGluPheAspValHisLeu 363
 :::::
 Db 31 TCTTCTTTAAGAAATGCTGCTAGTCATGTGTGGAGCCTTTGTAGAAATTTGACGTACACCTT 90

QY 364 SerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMetLysLys 383
 :::::
 Db 91 TCAAGAGGACTTTGTGCGCGGTGTATATCATGATCTTACCTGTTGTTGACTATGAAAAAG 150

QY 384 LysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAsp 403
 :::::
 Db 151 AAATTTGATGCTGATCCAGTTGAATTTATTTGAATTTCCAGTAAAGAAATTAACATTTGAC 210

QY 404 GlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGlu 423
 :::::
 Db 211 CAACCTCCAGTTGTTAAAGCTCACTCATGTGACTGCMCTGAAATCTTAAGGATCGGAAGAA 270

QY 424 SerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMet 443
 :::::
 Db 271 TCTGTGGTTTCAGAGGAAATTCCTTTTCAGAAAATCAGCCATTTCTCTTCTTAAGATG 330

QY 444 ValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTyrIleCysGln 463
 :::::
 Db 331 GTTTTAGAGTCTTTGCCAGAGATGTAGGGTTTAACATTGAATTAATAAATGATCTGCCAG 390

QY 464 GlnArgAspGlyMetTrrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeu 483
 :::::
 Db 391 CMAAGGGATGGAATGTGGATGGTAACCTTATCAACATATTTTGACATGAATCTGTTTTTG 450

QY 484 AspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSer 503
 :::::
 Db 451 GATATAATTTTMMAACTGTTTGTAGAAAATTTCTGGAGAGGAGAAATAGTGTTCCTCA 510

QY 504 PheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPhe 523
 :::::
 Db 511 TTTGATGCGATATTTGCACAATGTTCGGCAAAAGCAACAATATCCGATACTATTT 570

QY 524 LeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThr 543
 :::::
 Db 571 TTAACCTCAAGGAAAATCTGAGATTTATCTCTGAACCTCATGGACCTCAGATCTCGGACAACC 630

QY 544 ProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGlu 563
 :::::
 Db 631 CCCATTGCAATGAGCTTTTGCAAGTTTGAATAATCTACTGGGGATAAATGTACTACTGAA 690

QY 564 AspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePhe 583
 :::::
 Db 691 GACTTGCTCAAGAACCCATCTCTATATCAAGAGGCAAAAGTAAGGACTAGTCAATATTC 750

QY 584 CysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyVal 603
 :::::
 Db 751 TGCTGGGGTGATGATACCAATGATCTCGAAAACAGAGAAATTTGAAGAACTTGGAGTT 810

QY 604 AsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePheGln 623
 :::::
 Db 811 AATGGTCTAATTTATGATAGGATATATGATTTGGATGCTGCAACCAACCAATATATTCNA 870

QY 624 ValGluGlnLeuGluArgLysGlnGluLeuProGluLeuLysSerCysLeuCysPro 643
 :::::
 Db 871 GTGAGCAWTTGGAAACGCTGAAAGCAAGAAITGCAGAGCTTAAGAGCTGTTTGTTCOM 930

QY 644 ThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHisValAspAla 663
 :::::
 Db 931 CT-GTTAGCGCTTTGTTCCYCA-TCCTTTGTGGGAGTGTGATATCMT-GTGGATGCC 987

RESULT 9
 BM478430 1128 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT 6508010 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5578502
 DEFINITION 5', mRNA sequence.

ACCESSION BM478430
 VERSION BM478430.1 GI:18527472
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1128)

REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LHAM12333 row: k column: 15
High quality sequence stop: 618.
Location/Qualifiers
1. 1128
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5578502"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dn.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 3 47e-158 Length: 1128
Score: 1486.50 Matches: 324
Percent Similarity: 89.15% Conservative: 13
Best Local Similarity: 85.71% Mismatches: 26
Query Match: 42.03% Indels: 15
DB: 4 Gaps: 3

US-10-047-855-3 (1-672) x BM478430 (1-1128)

QY 217 TyrSerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGlu 236
DB 3 TTCGGTCGAGACGATGGAACCAATTAACCTGGAATTAATCTTTGATTTTTCGAAGA 62
QY 237 AspLeuSerGluHisValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCys 256
DB 63 GATCTCAGTGACGATGATTTGAGGTGATGCTTCTGACATGATGTTGAGTACAGCTGT 122
QY 257 LeuLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIle 276
DB 123 CTCCTTATCATCCACATTTCTGAGAGTGGAAGAGTGTGGAATCTTACTTCTCCCATC 182
QY 277 MetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleLeuLys 296
DB 183 ATGACGAGAAATCCCGGAAACAATAGCAAGTGAAGTGTGACTATATAATTAATTAAG 242
QY 297 ProLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArg 316
DB 243 CCATTACCAAGGATACAGTTGTGACATGAATCTTCAATTTTCAAGTATTTGAAGCAAGA 302
QY 317 IleProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAla 336
DB 303 ATACCAATTGGATTTGGCCATCGAGTGCAGGAACCTCTCAACCACTGCCAGTGGCT 362
QY 337 LysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPhe 356
DB 363 AAGTTTCAGAAAAATACATTTCTTTTAAGAAATGCTGCTAGTCAATGCTGAGCCCTT 422
QY 357 ValGluPheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThr 376
DB 423 GTAGAATTTGACGTACACCTTTCAAGGACTTTGTGCCGCTGTATATCATGATCTTACC 482
QY 377 CysCysLeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIlePro 396
DB 483 TGTGTGTGATGATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTGAAATTTCA 542
QY 397 ValLysGluLeuThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeu 416
DB 543 GTAAAGAAATTAACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTG 602
QY 417 LysSerLysAspArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGln 436

DB 603 AAATCTAAGGATCGAAAGAATCTGTGTTTCTGAGAGGAAATTCCTTTTCAGAAATCAG 662
QY 437 ProPheProSerLeuLysMetValLeuGluSer-LeuProGluAspVal-GlyPheAsnI 456
DB 663 CCAATTCCTTCTCTTAAGATGTTTGTAGTCTTTTGCAGAGATGTGATGTTTAAACA 722
QY 456 LeGluIleLysTrpIleCysGlnGlnArgAspGlyMet-TrpAspGlyAsnLeuSerThr 475
DB 723 TTGAATTAATTAATGATCTGCCAGCAAGGATGGAATGGGGGATGGTAATCTTATTCACA 782
QY 476 TyrPheAspMetAsnLeuPheLeuAspIleLeuLysThrValLeuGluAsnSerGly 495
DB 783 TATTTTGCATGATCTGTTTGTGATATAATTTTAAAAACTGTTTTTAAAAATCTCGG 842
QY 496 ---LysArgArgIleValPheSerSerPheAspAlaAsp-IleCysThrMet-ValArgG 514
DB 843 GAGAAGGAGAAAAAGGTTTTTCTTCATTTGATGAGATTAATTTGCCCACTGGTTCGGC 902
QY 514 InLysGlnAsnLysTyrPro---IleLeuPheLeuThrGlnGlyLysSerGlu-IleTyr 532
DB 903 AAAGCGAAACAATTAATCCCATCTAATTTTACTCCAGGAAATCTGAAAAATTTAT 962
QY 533 ProGluLeu-MetAspLeu-ArgSerArgThrThrProIleAlaMet-SerPheAlaGln 551
DB 963 CCGGACTCCATGGAATTCAGACTCGGACAAACCCCAATTCGAATGACCTTTTGGCCG 1022
QY 552 PheGluAsnLeu---LeuGlyIleAsnValHisThr--GluAspLeuLeuArgAsn-Pro 569
DB 1023 TTTGAAAAATTCACCTGGGGGATTAAGTACCTACTCTGAGAACTTGCTCAAAACCCCT 1082
QY 570 SerTyrIleGlnGlnAlaLysAlaLysGly 579
DB 1083 TCCCAATTTCAAGCAAAAGTCTAAGGA 1112
RESULT 10
LOCUS CB196769 901 bp mRNA linear EST 05-FEB-2003
DEFINITION AGENCOURT 11258308 NIH MGC 135 Mus musculus cDNA clone
IMAGE:30135584 5', mRNA sequence.
CB196769
CB196769.1 GI:28224876
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 901)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: NDAM0033 row: n column: 09
High quality sequence stop: 660.
Location/Qualifiers
1. 901
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30135584"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_135"
/note="vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5, 13.5,

14.5, and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. cDNA
enrichment: >1k bp. Average insert size 1.6k bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'-GACTAGTCTAGATCGAGCGCGCC(T)3'. Tissue contributed
by, David Rowe. Library constructed by ResGen, Invitrogen
Corp."

ORIGIN

Alignment Scores:

Pred. No.: 1-59e-144 Length: 901
Score: 1365.50 Matches: 276
Percent Similarity: 95.6% Conservative: 11
Best Local Similarity: 92.00% Mismatches: 5
Query Match: 38.61% Indels: 9
DB: 6 Gaps: 1

US-10-047-855-3 (1-672) x CB196769 (1-901)

QY 102 LeuGluSerGluIleIleAspGlyGlnPheGlyIleHisAsnGlyValGluThr 121
DB 12 CTAGAAAGTGAATCATTTACGATGACGATGTTGGCATCCACAAATGGTTCGAACA 71
QY 122 LeuAspSerGlyTyrLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerGlu 141
DB 72 CTGGATTCTGGATGCTTACATGTGAGCTGAATAAGATTCGCTCGCATTTTCTGAG 131
QY 142 LysProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLysLeu 161
DB 132 AAACCTCTGTTCAATTAGCAAGAAAAGTTCAAAAATCGAGATTTAGGTTAAAGCTC 191
QY 162 ThrLeuGluGlyLeuGlu-----AspAspAspAspArgValSerProThrVal 178
DB 192 ACACCTCGAGGCTCTGGAGGAAGATGATGATGATGATGATGATGATGATGATGATG 251
QY 179 LeuHisLysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCys 198
DB 252 CTTTCAAAATGTCACAGCCTCGAGATATCTTAAATGAAGTGAATGAGTTCAGTGC 311
QY 199 ArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTyrThrGluTyrSer 218
DB 312 AGGCACTCACGCCAAGATGGGTATGGCTTACAGCCGATCGTTGACAGAGTACAGC 371
QY 219 IleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeu 238
DB 372 ATACAGACAAATGGAACCAAGATAATCTGGAGCTCATCTTTGACTTTTGGAGGAATCTC 431
QY 239 SerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeu 258
DB 432 AGTGAGCATGTAGTTGAGGTGATGCTTCTCTGGACACGTTGGCACAGCATGCTCCTG 491
QY 259 SerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSer 278
DB 492 TCTTACATGCTGAGAGTGGGAAGCGTGGAACTTACTCTTCCCATCATGAGC 551
QY 279 ArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeu 298
DB 552 AGAAATTCAGAAAATATAGCAAGTCAGAGTTGATTTATCATCATCAAGCCATTA 611
QY 299 ProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTyrPheProArgIlePro 318
DB 612 CCTGATATAGTGTGTCTATGAGCTCTTCAATTTTCCAAAGTATTTGAAACCAAGATACCA 671
QY 319 LeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGluLeuAlaLysVal 338
DB 672 TTGGACGTTGGACATCGTGGTGGCANGGAACCTCAACAGCACTGCCAGCTAGCTAAAGTA 731
QY 339 GlnGluAsnThrIleAlaSerLeuArgAsnAlaLaSerHisGlyValAlaPheValGlu 358
DB 732 CAGGAATATCTATCGCTTCTTTAAA-AATGCTGCCAGTCAATGGCGCAGCATTTGTAGAA 790
QY 359 PheAspValHisLeu-SerLysAsp-PheValProValValTyrHisAsp-LeuThrCys 377

DB 791 TTTGATGTCCACCTTTTCAAAGGACTTTTGTGCCCGGTGTATCATGAACTCACCTGC 850
QY 378 CysLeuThrMetLysLysPheAsp-AlaAspProVal-GluLeuPhe 393
DB 851 TGTCTGACCATGAAGGAATATGAAGCTGTATCCCGTTTGAATTTGTT 900

RESULT 11

BM457674

LOCUS

DEFINITION AGENCOURT 6411862 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5498838
5', mRNA sequence.

ACCESSION

BM457674.1 GI:18506714

VERSION

BM457674.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Straubeberg, Ph.D.

Email: Cgapbs-remail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM12131 row: d column: 07

High quality sequence stop: 652.

FEATURES

Location/Qualifiers

1..882

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="IMAGE:5498838"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_85"

/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.867 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1-43e-139 Length: 882

Score: 1322.00 Matches: 251

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 37.38% Indels: 0

DB: 4 Gaps: 0

US-10-047-855-3 (1-672) x BM457674 (1-882)

QY 422 LysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSerLeu 441

DB 1 AAAGAATCTGTGGTTCAGGAGGAAAAATTCCTTTTCAGAAAAATCAGCCATTTCTCTCT 60

QY 442 LysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIle 461

DB 61 AAGATGTTTGTAGAGTCTTTGCCAGAGAGATGTAGGGTTTAACTTGAATAAATGGATC 120

QY 462 CysGlnGlnArgAspGlyMetTyrAspGlyAsnLeuSerThrTyrPheAspMetAsnLeu 481

DB 121 TGCAGCAAGGAGTGAATGTGGATGGTAACTTATCAACATATTTGACATGAATCTG 180

QY 482 PheLeuAspIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPhe 501

```
Db 181 TTTTGGTATATTTTAAACTGTTTTAGAAAAATTCGGAAGAGAGAGATAGTGT 240
Qy 502 SerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIle 521
Db 241 TCTTCAATTTGATGACAGATATTTGCACAAATGTTGCGCAAAAGCAGACAAATATCCGATA 300
Qy 522 LeuPheLeuThrGlnGlySerGluIleTyrProGluLeuMetAspLeuArgSerArg 541
Db 301 CTATTTTAACTCAAGGAAATCTGAGATTTATCCTGAACATCAATGGACCTCAGATCTCGG 360
Qy 542 ThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHis 561
Db 361 ACAACCCCATTCGATGAGCTTTGCACAGTTTGAAATCTACTGGGGATAAATGTACAT 420
Qy 562 ThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuVal 581
Db 421 ACTGAAGACTTGTCTCAGAAACCCATCCTATATTCAGAGGCGAAAGCTAAGGGACTAGTC 480
Qy 582 IlePheCysTyrGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeu 601
Db 481 ATATTTCTGCTGGGGTGATGATACCAATGATCCTGAAACACAGAGGAATTTGAAGGAACCT 540
Qy 602 GlyValAsnGlyLeuIleTyrAspArgIleTyrAspTyrMetProGluGlnProAsnIle 621
Db 541 GGAGTTAATGGTCTTAATTTATGATAGGATATATGATTTGGATGCTTGACACCAATATA 600
Qy 622 PheGlnValGluGlnLeuGluArgLysGlnGluLeuProGluLeuLysSerCysLeu 641
Db 601 TTCCAAGTGGAGCAATTTGGAACGCTGAAGCAGGAATTTGCCAGAGCTTAAGAGCTGT 660
Qy 642 CysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHisVal 661
Db 661 TGTCCCACTGTGTAGCGCTGTGTTCCTCATCTTTGTGTGGGAGTCTGATATCATGTG 720
Qy 662 AspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 721 GATGCCAAGCGCATTTGATAACGTGGAGATGCT 753

RESULT 12
BX403389 818 bp mRNA linear EST 28-APR-2004
LOCUS BX403389 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1059YA01 5-PRIME, mRNA sequence.
ACCESSION BX403389
VERSION BX403389.2 GI:46846825
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 818)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30615152.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9838.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS5AA0152D05RMI&c=9838.r.
Location/Qualifiers
1. .818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1059YA01"
```

FEATURES
source

```
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/Note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Alignment Scores: 1.64e-139 Length: 818
Pred. No.: 1321.00 Matches: 261
Score: 96.3% Conservative: 3
Percent Similarity: 96.3% Mismatches: 7
Best Local Similarity: 95.26% Indels: 3
Query Match: 37.35% Gaps: 0
DB: 5

US-10-047-855-3 (1-672) x BX403389 (1-818)

Qy 345 SerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAspValHisLeuSer 364
Db 3 TCTTTAAGAAATGCTGCTAGTCAATGTCAGCCTTTGTAGAAATTTGACGTACACCTTTCA 62
Qy 365 LysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMetLysLysLys 384
Db 63 AAGGACTTTGTGCCGTGGTATATCATGATCTTACCTGTGTGTGACTATGAAAAAGAAA 122
Qy 385 PheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAspGln 404
Db 123 TTTGATGCTGATCCAGATTGAATTTATTTGAAATTTCCAGTAAAGAAATTAACATTTGACCAA 182
Qy 405 LeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSer 424
Db 183 CTCACGTTGTTAAAGCTCACTCATGTGACTGCACATGAAATCTAAGGATCGGAAAGAACT 242
Qy 425 ValValGlnGluGlnAsnSerPheSerGluAsnGlnProPheProSerLeuLysMetVal 444
Db 243 GTGGTTTCAGGAGGAAAAATTCCTTTTTCAGAAAAATTCAGCCATTTTCTTCTTAAGATGGTT 302
Qy 445 LeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTyrIleCysGlnGln 464
Db 303 TTAGAGTCTTTGCGAAGAGATGAGGGTTTAACTTTGAAATTAATGATGATCTGCCAGCAA 362
Qy 465 ArgAspGlyMetTyrAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeuAsp 484
Db 363 AGGATGGAATGTGGGATGGTAACTTATCAACATATTTTGACATGAATCTGTTTTGGAT 422
Qy 485 IleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSerPhe 504
Db 423 ATAATTTTAAAAACTGTTTTAGAAAAATTCCTGGGAGAGGAGATAGTGTTCCTTCATTT 482
Qy 505 AspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeu 524
Db 483 GATGCGATATTTGCGCAATGTTTGGGCAAAAGCAGAACAAATATCCGATCTATTTTAA 542
Qy 525 ThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThrPro 544
Db 543 ACTCAAGAAAAATCTGAGATTTATCTGAACTCATGCACTCAGATCTCGGCAACCCCC 602
Qy 545 IleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGluAsp 564
Db 603 ATTGCAATGAGCTTTGCACAGTTTTGAAATCTACTGGGGATAAATGTACATCTGAAAC 662
Qy 565 LeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCys 584
Db 663 TTGCTCAGAAACCCATCCTATATTTCAAGAGGCAAGCTAAGGACTAGTCATATTCG 722
Qy 585 TrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyValAs 604
Db 723 TGGGGTGATGATCCCAATGATTCCTGAAACAGAGAGAAATTTAGGAGACCTTGGAGTTAA 782
Qy 604 n-GlyLeuIleTyrAspArgIleTyrAsp-TripMet 615
Db 783 TGGGCTTAATTTTGGATAGGAAATATGATTTGGATG 818
```

RESULT 13
CD466056 783 bp mRNA linear EST 04-JUN-2003
LOCUS LeukN2_1_G06_g1_A024 Unstimulated peripheral blood leukocytes N2
DEFINITION Equus caballus cDNA clone LeukN2_1_G06_A024 5', mRNA sequence.
ACCESSION CD466056
VERSION 1
KEYWORDS EST.
SOURCE CD466056.1 GI:31387324
ORGANISM Equus caballus (horse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE 1 (bases 1 to 783)
AUTHORS Vandenplas, M., Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S.,
Moore, J. N., Liang, C., Sun, F., Sullivan, R., Shah, M. and Pratt, L. H.
TITLE An EST database from equine (Equus caballus) unstimulated
peripheral blood leukocytes
JOURNAL Unpublished (2003)
COMMENT Other ESTs: LeukN2_1_G06_b1_A024
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 593 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; tissue and RNA were prepared in the Department of Large
Animal Medicine, University of Georgia; sequencing done in the
Laboratory for Genomics and Bioinformatics, University of Georgia.
Sequence ends have been trimmed to exclude vector and regions below
phred quality 16. Three-prime sequences are presented as their
reverse complement and have been trimmed to exclude polyA.
Seq primer: Sugs (CTCTGCTCTAAAGCTGGC).

FEATURES

Location/Qualifiers
1..783
/organism="Equus caballus"
/mol_type="mRNA"
/strain="thoroughbred"
/db_xref="taxon:9796"
/clone="LeukN2_1_G06_A024"
/sex="male"
/tissue type="blood"
/cell type="leukocytes"
/lab_host="DH108-T1 phage-resistant E. coli"
/clone_lib="Unstimulated peripheral blood leukocytes N2"
/note="Organ: circulatory system; Vector: pME18S-FL3;
Site 1: XhoI; Site 2: XhoI; The library was prepared from
polyA+ RNA from unstimulated equine peripheral blood
leukocytes isolated from a healthy adult horse.
Double-stranded cDNA was cloned unidirectionally into
different DraIII sites of the pME18S-FL3 vector (5'-prime
DraIII site is CACTGTGTG, 3'-prime DraIII site is
CACCATTGTC). XhoI excises the cDNA insert."

ORIGIN

Alignment Scores:
Pred. No.: 2,276-139 Length: 783
Score: 1319.50 Matches: 257
Percent Similarity: 98.85% Conservative: 1
Best Local Similarity: 98.47% Mismatches: 2
Query Match: 37.31% Indels: 2
DB: 6 Gaps: 1

US-10-047-855-3 (1-672) x CD466056 (1-783)

QY 158 ArgValLysLeuThrLeuGluGlyLeuGluAspAspArgValSerProThr 177
DB 3 AGGGTAAGCTCACTAGACAGTCTGGAGGAGAGATGACATGATAGGCTGTCCACT 62
QY 178 ValLeuHisLysMetSerAsnSerLeuGluLeuSerLeuLeuSerAspAsnGluPhetLys 197

Db 63 GTTCTTCAAAAATGTCCACAGCCTGGAGATATCGTTAATAAGTGAACAATGAGTTCAAG 122
QY 198 CysArgHisSerGlnProGluCysGlyTyrrGlyLeuGlnProAspArgTrrThrGluTyrr 217
Db 123 TGCAGGCATTCACAGCCGAGTGGGTATGCTTACAGCCTGATCGCTGGACGGAGTAT 182
QY 218 SerTleGlnThrMetGluProAspAsnLeuGluLeuLeuPheAspPheGluGluasp 237
Db 183 AGCATACAGACAATGGAACCGGATAAAGCTGGAACCTGATATTCGATTTTTTGA---GAT 238
QY 238 LeuSerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeu 257
Db 239 CTCAGTGAACATAGTTCAGGGCGATGCCCTTCCTGACATGTTGGGTACAGCGTGCCTC 298
QY 258 LeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMet 277
Db 299 TTATCATCCACCATGCTGAGAGTGGGAGAGTCTGGATCTTACTCTTCCCACATCATG 358
QY 278 SerArgHisSerArgLysThrIleGlyLysValArgValAspTyrrIleIleLeuLysPro 297
Db 359 AGCAGAAATTCAGGAAACAAATAGGCAAGTGAAGTGTGACTATATAATTTAAGCCA 418
QY 298 LeuProGlyTyrrSerCysAspMetLysSerSerPheSerLysTyrrTrrLysProArgIle 317
Db 419 TTACCAGGATACAGTGTGACATGAAATCTTCATTTTCCAAAGTACTGGAAGCCAAAGAATA 478
QY 318 ProLeuAspValGlyHisArgGlyValGlyAsnSerThrThrAlaGlnLeuAlaLys 337
Db 479 CCATTGATGTTGGGCATCGAGGTGCAGGGAATTTCTACACAACTGCTCAGCTTGCTAAA 538
QY 338 ValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheVal 357
Db 539 GTTCAAGAAATACTATTGCTTCTTTAAGAAATGCTGCTAGTCATGCTGCGACGCTTTGTA 598
QY 358 GluPheAspValHisLeuSerLysAspPheValProValValTyrrHisAspLeuThrCys 377
Db 599 GAATTTGATGTACATCTTTCAAAAGATTTTGTGCTGGTGTATCATGATCTCACCTGT 658
QY 378 CysLeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProVal 397
Db 659 TGTTTGACTATGAAAGAAATTTGATGCTGATCAGTTGAAATTTTGAATTTCCAGTA 718
QY 398 LysGluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLys 417
Db 719 AAGGAATTAACATTTGACCAGCTCCAGTTGTTAAAGCTCGCTCATGTGACTGCACCTAAA 778
QY 418 Ser 418
Db 779 TCT 781

RESULT 14
BU503488 958 bp mRNA linear EST 12-SEP-2002
LOCUS AGENCOURT_8968328 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6491134
DEFINITION 5', mRNA sequence.
ACCESSION BU503488
VERSION BU503488.1 GI:22809721
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 958)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>
TITLE National Institutes of Health/
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1AM14042 row: a column: 23
High quality sequence stop: 597.

FEATURES

Location/Qualifiers
1..958

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clones="IMAGE:6491134"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1.44e-137 Length: 958
Score: 1305.00 Matches: 266
Percent Similarity: 89.44% Conservative: 22
Best Local Similarity: 82.61% Mismatches: 19
Query Match: 36.90% Indels: 17
DB: 5 Gaps: 4

US-10-047-855-3 (1-672) x BU503488 (1-958)

```
QY 169 AspAspAspArgValSerProThr-ValLeuHisLysMetSerAsnSerLeuGluIl 188
DB 1 GATGATGACGATAGGTTCTCTCCACTGGTTCTTCCACAAAATGTCACACGCTGGAGAT 60
QY 188 eSerLeuLeSerAspAsnGluPheLysCysArgHisSerGlnProGluCysGlyTyrGl 208
DB 61 ATCCCTTAATAGTGAACATAGTTCAAGTGCAGGCACTCAGCCAGAAATGTGGGTATGG 120
QY 208 YLeuGlnProAspArgTyrThrGluTyrSerileGlnThrMetGluProAspAsnLeuGl 228
DB 121 CTTACAGCCCGATCGTTGGACAGAGTACAGCATACAGACATGGAACACAGATAATCTGGA 180
QY 228 uLeuLePheAspPheGluGluAspLeuSerGluHisValValGlnGlyAspAlaLe 248
DB 181 GCTCATCTTTGACTTTTITAGCAAGATCTCAGTCAGCATGTAGTTCAGGGTGATGTTCT 240
QY 248 uProGlyHisValGlyThrAlaCysLeuLeuSerSerThrThrileAlaGluSerGlyLysE 268
DB 241 TCGTGACACGCTGGGCACAGCATGCTCTGTTCTTACCATGTCTGAGAGTGGGAAG 300
QY 268 rAlaGlyIleLeuThrLeuProIleMetSerArgAsnSerArgLysThrileGlyLysVa 288
DB 301 CGCTGGAATCCTTACTCTTCCCATCATGACGAGAAATTCAGAAAAACTATAGGCANAAGT 360
QY 288 lArgValAspTyrIleIleLysProLeuProGlyTyrSerCysAspMetLysSerSe 308
DB 361 CAGAGTTGATTTATCATCATCAAGGCAATACCTCGATATAGTGTCTTATGACGTCTTC 420
QY 308 rPheSerLysTyrTrpLysProArgIleProLeuAspValGlyHisArgGlyAlaGlyAs 328
DB 421 ATTTTCCAAGTATTTGAAACCAAGATACCATTTGGACGTTGGACATCGGTGGCAGGGAA 480
QY 328 nSerThrThrThrAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAs 348
DB 481 CTCACAAACGACTGCCCAAGTACTAAAGTACAGGAAATATATTCGCTCTTTTAGAAA 540
QY 348 nAlaAlaSerHisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheVa 368
DB 541 TGTGCGCAGTCATGGCGCAGCATTTGTAGAAATTTGATGTCACCTTTCAAGGACCTTGT 600
QY 368 lProValValTyrHisAspLeuThrCysCysLeuThrMetLysLeuLysPheAspAlaAs 388
```

Alignment Scores:

```
DB 601 GCCCGTGGTATCATGACCTCACCTGCTGTCTGACCATGAGGAATAATATGAGCTGA 660
QY 388 pProValGluLeuPheGluIleProValLysGluLeuThrPheAspGlnLeuGlnLeuLe 408
DB 661 TCCAGTTGAATGTTTGAATCCAGTAAAGAAATACCAITTTGACCAATCCAGTATT 720
QY 408 uLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSerValValGlnGl 428
DB 721 GAAGCTTTCTCATGTGACTGG-CTTAAAAAACCAAGACCGGGAACATCTTTGTATGAGA 779
QY 428 uGluAsn-SerPheSerGluAsnGlnProPheProSer-LeuLysMetValLeuGluSer 447
DB 780 GGAATAATTTTCTTTCTGAAAATCAGCCATTTCTCTCTAAAGAGGGTTTATGAATC 839
QY 448 LeuProGlu---AspValGlyPheAsn---IleGluIleLysTyrIleCysGlnGlnArg 465
DB 840 ATTTCACAGAAAATGTAGATTTTCTAATTGAATTACATGGA-ATTGCCCCACCC 898
QY 466 AspGly-----MetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPhe 482
DB 899 CAGGGAATGGAATAATAATATGGATGGC-----CCCAACTTATTTC 937
```

RESULT 15

BU145581 948 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT 8125917 Lupski dorsal root ganglion Homo sapiens cDNA
DEFINITION clone IMAGE:6177333 5', mRNA sequence.

ACCESSION

BU145581

VERSION

BU145581.1 GI:22659113

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 948)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabos-fo@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1AM13555 row: b column: 22
High quality sequence stop: 653.

FEATURES

Location/Qualifiers

1..948

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6177333"

/sex="male"

/tissue_type="dorsal root ganglia"

/dev_stage="adult, 36 yr"

/lab_host="DH10B"

/clone_lib="Lupski dorsal root ganglion"

/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:

NotI; Site_2: SalI; cDNA made by oligo-dr priming.

Directionally cloned using the following adaptors:

5'-TCGACCCACGCGCCG-3' and

5'-GACTAGTTCTAGATCGGAGCGCGCCCTT(15)-3'. Size selected >

1 kb for average insert length 1.7 kb. This is a primary

library, non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

College of Medicine) and is available through Life

Technologies."

Pred. No.: 2,72e-135 Length: 948
Score: 1285.00 Matches: 243
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 36.33% Indels: 0
DB: 5 Gaps: 0

US-10-047-855-3 (1-672) x BUI45581 (1-948)

```
QY 430 AsnSerPheSerGluAsnGlnProPheProSerLeuLysMetValLeuGluSerLeuPro 449
DB 3 AATTCCTTTTCAGAAATCAGCAATTCCTCTCTTAAGATGTTTATAGATCTTTGGCA 62
QY 450 GluAspValGlyPheAsnIleGluIleLysTrpIleCysGlnGlnArgAspGlyMetTrp 469
DB 63 GAAGATGTAGGGTTTAAACATTGAAATAAAATGGAATCTGCCAGCAAAAGGATGGAATGTGG 122
QY 470 AspGlyAsnLeuSerThrTyrrPheAspMetAsnLeuPheLeuAspIleIleLeuLysThr 489
DB 123 GATGGTAACCTTATCAACATATTTTGACATGAATCTGTTTTTTCGATATATATTTTAAAAACT 182
QY 490 ValLeuGluAsnSerGlyLysArgGileValPheSerSerPheAspAlaAspIleCys 509
DB 183 GTTTTAGAAATCTGGGAGAGGAGAAATAGTGTGTTTCTTCATTTGATGCAGATATTTGC 242
QY 510 ThrMetValArgGlnLysGlnAsnLysTyrrProIleLeuPheLeuThrGlnGlyLysSer 529
DB 243 ACAATGGTTTCGCAAAAGCAGAAATAATATCCGATACTATTTTAACTCAAGGAAAAATCT 302
QY 530 GluIleTyrrProGluLeuMetAspLeuArgSerArgThrThrProIleAlaMetSerPhe 549
DB 303 GAGATTTATCTCGAACTCATGGACCTCAGATCTCGGACACCCCAITTCGAATGAGCTTT 362
QY 550 AlaGlnPheGluAsnLeuGlyIleAsnValHisThrGluAspLeuLeuArgAsnPro 569
DB 363 GCACAGTTTGAATAATCTACTGGGGATTAATGTACATCTGAAGACTTGTCTCAGAAACCA 422
QY 570 SerTyrrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpGlyAspAspThr 589
DB 423 TCCTATATTCAGAGGCAAAAGCTAAGGACTAGTCATATTCCTGCTGGGGTGATCATACC 482
QY 590 AsnAspProGluAsnArgArgLysLysLeuGlyValAsnGlyIleuIleTyrrAsp 609
DB 483 AATGATCTCTGAAAACAGAAAGGAAATTTGAAGGAACCTGGAGTTTAATGGTCTAATTTATGAT 542
QY 610 ArgIleTyrrAspTipMetProGluGlnProAsnIlePheGlnValGluGlnLeuGluArg 629
DB 543 AGGATATATGATTGGATGCTGACCAACCAATATATTTCCAAGTGGAGCAATTTGGAACGC 602
QY 630 LeuLysGlnGluLeuProGluLeuLysSerCysLeuCysProThrValSerArgPheVal 649
DB 603 CTGAAGCAGGAATTCAGAGCTTAAGAGCTGTTTGTGTGCCACTGTTAGCGCTTTGTT 662
QY 650 ProSerSerLeuCysGlyGluSerAspIleHisValAspAlaAsnGlyIleAspAsnVal 669
DB 663 CCTCATCTTTGTGTGGGAGTCTGATATCCATGTGGATGCCAACGGCAATTTGATAACGTG 722
QY 670 GluAsnAla 672
DB 723 GAGAATGCT 731
```

Search completed: July 3, 2005, 10:56:38
Job time : 5796 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 3, 2005, 07:07:40 ; Search time 929 Seconds
(without alignments)
3248.529 Million cell updates/sec

Title: US-10-047-855-3
Perfect score: 3537
Sequence: 1 MTPSOVAPEIRGTLPLGEVF.....LCGESDIHVANGIDNVENA 672

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 13051180 seqs, 2245448946 residues
Total number of hits satisfying chosen parameters: 26102360

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame_plus_p2n.model -DB=FASTA
-O=/cgn2_1/USPTO_spoil/US10047855/runat_01072005_155530_2351/app_query.fasta.1.839
-DB=Pending_Patents_NA_New -QPMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATTRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -MINLEN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=200 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10047855 @CNC 1 1 824 @runat_01072005_155530_2351 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New:
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq2:
3: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:
5: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq3:
10: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq4:
11: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq5:
12: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq:
13: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq2:
14: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq3:
15: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2965.5	83.8	3381	12	US-11-136-527-4005 Sequence 4005, Ap
2	1399	39.6	2393	12	US-11-136-527-4018 Sequence 4018, Ap
3	1005	28.4	2792	12	US-11-097-143-5114 Sequence 5114, Ap

C	4	922	26.1	5439	12	US-11-097-143-5113	Sequence 5113, Ap
	5	667	18.9	2392	12	US-11-097-143-27182	Sequence 27182, A
	6	667	18.9	2457	12	US-11-097-143-26897	Sequence 26897, A
	7	667	18.9	7162	12	US-11-097-143-26896	Sequence 26896, A
	8	667	18.9	9667	12	US-11-097-143-27181	Sequence 27181, A
	9	666	18.8	2299	12	US-11-097-143-19967	Sequence 19967, A
	10	666	18.8	4299	12	US-11-097-143-19966	Sequence 19966, A
	11	661	18.7	2220	12	US-11-097-143-40595	Sequence 40595, A
	12	661	18.7	4220	12	US-11-097-143-40594	Sequence 40594, A
	13	634.5	17.9	693	1	PCT-US05-00517-1279	Sequence 1279, Ap
	14	550	15.5	2100	12	US-11-097-143-33686	Sequence 33686, A
	15	550	15.5	4100	12	US-11-097-143-33685	Sequence 33685, A
	16	494	14.0	3657	9	US-10-932-182A-2598	Sequence 2598, Ap
	17	476.5	13.5	1272	9	US-10-932-182A-81481	Sequence 81481, A
	18	466.5	13.2	757	15	US-60-669-241-22115	Sequence 22115, A
	19	292.5	8.3	2357	15	US-60-655-875-78279	Sequence 78279, A
C	20	292.5	8.3	3176	15	US-60-655-875-2174	Sequence 2174, Ap
C	21	287.5	8.1	2849	12	US-11-097-143-6409	Sequence 6409, Ap
	22	231.5	6.5	531	15	US-60-655-875-93796	Sequence 93796, A
	23	224	6.3	882	9	US-10-703-032-3896	Sequence 3896, Ap
	24	222	6.3	781	9	US-10-703-032-16655	Sequence 16655, A
C	25	199	5.6	914	15	US-60-655-875-2175	Sequence 2175, Ap
	26	197	5.6	634	15	US-60-655-875-62753	Sequence 62753, A
	27	197	5.6	1116	15	US-60-655-875-32447	Sequence 32447, A
	28	195	5.5	707	15	US-60-669-241-5589	Sequence 5589, Ap
C	29	186	5.3	4001	15	US-60-655-875-10470	Sequence 10470, A
	30	186	5.3	4001	15	US-60-655-875-73678	Sequence 73678, A
	31	180	5.1	736	9	US-10-703-032-14998	Sequence 14998, A
	32	178	5.0	2703	9	US-10-932-182A-80085	Sequence 80085, A
C	33	174	4.9	2089378	9	US-10-526-324-1	Sequence 1, Appli
	34	174	4.9	2089378	9	US-10-526-324-342	Sequence 342, App
C	35	174	4.9	2089378	9	US-10-526-324-723	Sequence 723, App
	36	174	4.9	2089378	9	US-10-526-324-1087	Sequence 1087, Ap
	37	174	4.9	2089378	9	US-10-526-324-1469	Sequence 1469, Ap
	38	174	4.9	2089378	9	US-10-526-324-1838	Sequence 1838, Ap
	39	172.5	4.9	4830849	2	PCT-IB03-06509-1355	Sequence 1355, Ap
C	40	169.5	4.8	605	9	US-10-703-032-34240	Sequence 34240, A
	41	159.5	4.5	636	15	US-60-655-875-88995	Sequence 88995, A
	42	150	4.2	481	15	US-60-655-875-96527	Sequence 96527, A
	43	148.5	4.2	2615	15	US-60-655-875-74699	Sequence 74699, A
	44	148.5	4.2	2848	15	US-60-655-875-11318	Sequence 11318, A
	45	145	4.1	14570	13	US-11-031-175-1012	Sequence 1012, Ap

ALIGNMENTS

RESULT 1
US-11-136-527-4005
; Sequence 4005, Application US/11136527
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4005
; LENGTH: 3381
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4005

Alignment Scores:
Pred. No.: 1.04e-263
Score: 2965.50
Percent Similarity: 90.76%
Best Local Similarity: 85.10%
Query Match: 83.84%
DB: 12
Length: 3381
Matches: 571
Conservative: 38
Mismatch: 33
Indels: 29
Gaps: 6

US-10-047-855-3 (1-672) x US-11-136-527-4005 (1-3381)

QY 1 MetThrProSerGlnValAlaPheGluLeuArgGlyThrLeuLeuProGlyGluValPhe 20
DB 186 ATGACACCTTCTCAGGTCACTTTGAAATAGAGGAACCTTTTACCAGGAGAGGTCTTT 245
QY 21 AlalleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
DB 246 GCAATGTTGTGAAACTGTGATGCCTTGGGAACTGGAGTCTCAAAATGCTGTGCCT--- 302
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
DB 303 CTTACTGAGAGTGACACAGGCGAAAGT---GTATGGAAGAGCAGTATTGTTCTTAGTAGA 359
QY 61 GlyValSerValGlnTrpArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
DB 360 GGAATGTCGTCGTAAGTACCGTACTTACAGGCTCTTTTAGAACCAAGACATATCGGT 419
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
DB 420 GGTCCATGTCAAGTATAGTTCACAAAGTGGAGACTCATCTACACCAAGCATCAATAACC 479
QY 101 ProLeuGluSerGluIleIleLeuAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
DB 480 CCTTTAGAAAACGAAATCATATTACCATGGACAATTTGGAATCCAATGGTGTGAA 539
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
DB 540 ACATGGATCTCGATGGATGGCTTACCTGTACACTGAAATAGACTGCGTCTGCATTTTCT 599
QY 141 GluLysProProValSerIleThrLysLysLeuLysLysLeuLysLysLeuLysLys 160
DB 600 GAGAACTCTCTGTTTCAATTACCAAGAAAAGTTCAAAAATCTAGATTTAGGGTAAAG 659
QY 161 LeuThrLeuGluGluLeuGlu-----AspAspAspAspArgValSerProThrVal 178
DB 660 CTTACACTAGAGGTCTGGAGGAAGATGATGACGACGATGATAGGCAATCTCCCACTGTT 719
QY 179 LeuHisLysMetSerAsnSerLeuGluIleSerLeuLysLeuAspAsnGluPheLysCys 198
DB 720 CTTCAAGATGTCCTAGCTGGAGATATCTTATAGTACAAATGAGTTCAGTGC 779
QY 199 ArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSer 218
DB 780 AGGCATCTCAGCCAGAGATGGGTATGGCTTACAGCTGACCGCTGGACAGATACAGC 839
QY 219 IleGlnThrMetGluProAspAsnLeuGluIlePheAspPheGluGluAspLeu 238
DB 840 ATACAGACAATGGAGCGGACCACTTGAACCTCATCTTTTGTGAGGAAGATCTC 899
QY 239 SerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeu 258
DB 900 AGTGAGCATGTATGTCAGGGTGTATGTTCTTCCGACATGTGGGACAGCATGCCTCCTG 959
QY 259 SerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSer 278
DB 960 TCAITACCAATTGCTGAGTGTGAGAGGCGCTGGAATCCTTACTCTTCCATCATGAGC 1019
QY 279 ArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeu 298
DB 1020 AGAAGTCCAGAAAACTATAGGCAAGTCAGAGTTGATTTTATCATCATCAAGCCATTA 1079
QY 299 ProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIlePro 318
DB 1080 CCAGATATAGTTGTTCTATGCGAGTCTTCATTTCTTCCAAAGTATTGGAAACCAAGATACCA 1139
QY 319 LeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrThrAlaGlnLeuAlaLysVal 338
DB 1140 CTGATGTTGGACATCTGTGTGAGGGAACCTCAACAACACTGCCAGCTGGCTAAAGTA 1199
QY 339 GlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGlu 358

DB 1200 CAGGAAATACATATTGCTTCTTTAAGAAATGCTGCCAGCCATGGTGCAGCATTTTGTGGAA 1259
QY 359 PheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCys 378
DB 1260 TTTGATGTCCACTTTCAGAGGACTTAGTGCCTGTAGTGTATCATGATCTCACCTGCTGT 1319
QY 379 LeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLys 398
DB 1320 TTAACCTATGAAAGGAAATATGAAGCTGATCCAGTTGAATGTTTGAATCCAGTAAAG 1379
QY 399 GluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSer 418
DB 1380 GAATTAACATTCGACCACTCCAGTTATTGAAAGCTTTCTCATGTGACTGCATTAANAACC 1439
QY 419 LysAspArgLysGluSerValGlnGluGluAsnSerPheSerGluAsnGlnProPhe 438
DB 1440 AAAGACCAACAAATGATGCTGAGGAGGAAATTCCTTTCTGAAACCAACCATTT 1499
QY 439 ProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIle 458
DB 1500 CTTTCTTTAAGATGGTTTAGAGTCATTGCCAGAAAATGTAGGATTTAATATAGAATA 1559
QY 459 LysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAsp 478
DB 1560 AAATGGATTTGCCAACACACAGGAGTGGAGTATGGCGCGCAACTTATCGACATATTTTGAT 1619
QY 479 MetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArg 498
DB 1620 ATGAATGCATTTTGGATATTAATTTTAAAACTGTTTATAGAAATTCGCGGAAGAGAGA 1679
QY 499 IleValPheSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLys 518
DB 1680 ATAGTATTTTCTTCAITTTGATGACAGATCTGTACAATGTTGCGCAGAAAACAAA 1739
QY 519 TyrProIleLeuPheLeuThrGlnGlySerGluIleTyrProGluLeuMetAspLeu 538
DB 1740 TATCCCATATATTATTTGACCAAGGAAGTCTGACATTTTACCTGAACCTCATGGACCTC 1799
QY 539 ArgSerArgThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIle 558
DB 1800 AGATCTCGGACAAACACCATTTGCAATGAGCTTTGCACAGTTTGAAAATATTTGGGATA 1859
QY 559 AsnValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLys 578
DB 1860 AATGCCATCTAGGATCTCTTAGAAACCATCTTATGTCCAAGAGGCAAAAGATAAG 1919
QY 579 GlyLeuValIlePheCysTrpGlyAspThrAsnAspProGluAsnArgArgLysLeu 598
DB 1920 GGATTTGTCATATTCTGCTGGGTGATGATACCAATGATCCTGAAAACAGAAAGAACTG 1979
QY 599 LysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGln 618
DB 1980 AAGGAATTTGGAGTAAATGGTCTAATATATATAGTAT----- 2018
QY 619 ProAsnIlePheGlnValGluGlnLeuGluArgLeuLysGlnLeu----- 634
DB 2019 -----TTGTTTTTGTAAAAATCTCATGGAATGTTTCAACAGTGTAGTTTATCTA 2072
QY 635 -----ProGluLeuLysSer-----CysLeuCys 642
DB 2073 TTTTAACTATTTTAAATTAGATAGTTTAGCCTAAAGTTTATCTTGTGACTGTGACCTT 2132
QY 643 ProThrValSerArgPheValProSerSerLeu 653
DB 2133 TCCAGGTGTTGAGATATGTCAAAAGCCACTTA 2165

RESULT 2

US-11-136-527-4018
; Sequence 4018, Application US/11136527
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes

```
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4018
; LENGTH: 2393
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-4018

Alignment Scores:
Pred. No.: 8,67e-119 Length: 2393
Score: 1399.00 Matches: 261
Percent Similarity: 95.8% Conservative: 16
Best Local Similarity: 90.31% Mismatches: 12
Query Match: 39.55% Indels: 0
DB: 12 Gaps: 0

US-10-047-855-3 (1-672) x US-11-136-527-4018 (1-2393)

QY 384 LysPheAspAlaAspProValGluLeuPheGluLeuProValLysGluLeuThrPheAsp 403
DB 71 AAATATGAAGCTGATCCAGTTGAATGTTGAAATCCCAAGTAAGGAATTAACATTCGAC 130
QY 404 GlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGlu 423
DB 131 CMACTCCAGTTAATGAGCTTTCTCATGTGACTGCACTAAACCAAGACCAAGCAACAA 190
QY 424 SerValGlnGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMet 443
DB 191 TGTATGGCTGAGGAGGAAATCTCTTCTGAAACCAACCACTTCTCTCTTAAGATG 250
QY 444 ValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGln 463
DB 251 GTTTTAGAGTCTATCCCAAGAAATAGGATTTAATATAGAAATAAATGGAATTTGCCAA 310
QY 464 GlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeu 483
DB 311 CACAGGATGGATGGACGGCACTTATCGACATATTTTGATGATGATGATGATGATG 370
QY 484 AspileLeuLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSer 503
DB 371 GATATAATTTAAAACTGTTTGAAGAAATTCGGGAGAGAGAGAGATAGTATTTCTTCA 430
QY 504 PheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPhe 523
DB 431 TTTGATGCAGACATCTGTACATGTTTCGGCAGAAACAAACAAATATCCCATATATTT 490
QY 524 LeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThr 543
DB 491 TTGACCCCAAGGAAAGTCTGACATTTACCTCGAACTCATGGACCTCAGATCTCGGACACA 550
QY 544 ProfileAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGlu 563
DB 551 CCCATTGCAATGAGCTTTGCAGATTTGAAATATTTGGGATTAATGCCCATACTGNA 610
QY 564 AspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePhe 583
DB 611 GATCTCTTAGAAACCCATCTATGTCCAAGAGGCAAAAGATAAGGATTTGGTCATATTC 670
QY 584 CysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuVal 603
DB 671 TGCTGGGTGATGATACCAATGATCTGAAACACAGAGAAACTGAAGAAATTTTCGAGTA 730
QY 604 AsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePheGln 623
DB 731 AATGCTCTAATATATGATAGGATATACGATTCGATTCGCTGACACCAACCAATATATCCA 790
QY 624 ValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCysLeuCysPro 643
DB 939 HisLeuGlnProArgSerIleThrProLeuGluSerGluIleIleIleAspGlyGln 112

Db 791 GTGGAGCAACTGGAGCGCTGAAGCGAGAAATTGCCAGAGCTTAAGAACTGTTGTGTCCC 850
QY 644 ThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHisValAspAla 663
DB 851 ACTGTTAGCCACTTCATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 910
QY 664 AsnGlyIleAspAsnValGluAsnAla 672
DB 911 AACGCGATTGATAATGTGGAACGCT 937

RESULT 3
US-11-097-143-5114
; Sequence 5114, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5114
; LENGTH: 2792
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-5114

Alignment Scores:
Pred. No.: 3,4e-82 Length: 2792
Score: 1005.00 Matches: 250
Percent Similarity: 52.29% Conservative: 116
Best Local Similarity: 35.71% Mismatches: 244
Query Match: 28.41% Indels: 91
DB: 12 Gaps: 22

US-10-047-855-3 (1-672) x US-11-097-143-5114 (1-2792)

QY 14 LeuLeuProGlyGluValPheAlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsn 33
DB 529 CTGGCGCCCAACAGACAGCTTGCCCATAGTGGTAAATTCGGAGCGCTGGGCAACTGGCAG 588
QY 34 ProGlnAsnAlaValAlaLeuLeuPro---GluAsnAspThrGlyGluSerMetLeuTrp 52
DB 589 CACTCGGAGCGGCTATTTTGTCCAAAATGAGGAGGACAAATGAGTAGTAGTATCTGTGG 648
QY 53 LysAlaThrIleValLeuSerArgGlyValSerValGlnTyrArgTyrPheLysGlyTyr 72
DB 649 ACAGGAGAGATCTATATTCGCGCTCACTGTGCACACGGAATATCGCTATATGTC---TGT 705
QY 73 PheLeuGluProLysThrIleGlyGlyProCysGlnValIleValHisLysTrpGluThr 92
DB 706 GCCGTCGATCCC-----GGCAGGAGCAGCTGATTTGTCGCGGTTGGGAAACG 753
QY 93 HisLeuGlnProArgSerIleThrProLeuGluSerGluIleIleIleAspGlyGln 112
```



```
QY 362 HisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMet 381
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1355 CATTTGACTGCTGATGGTGTGCTGTGATTTATCAGCATTTCCGAGTCCGAACTGCTCCG 1414

QY 382 LysLeuLysPheAspAlaAspProValGluLeuPheGluLeuProValLysGluLeuThr 401
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1415 CTGTGGCAAGCAGATCAGCCGCGCAGACCTGAGTGTGTGATCAAGACATAAAGC 1474

QY 402 PheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArg 421
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1475 TATGACTGCTCAAAAGGCTGGCATCTTCTGTGTCGCA----- 1516

QY 422 LysGluSerValValGlnGluGluAenSerPheSer-----GluAenGln 436
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1517 -----GGCCAAGTGAGGAGTATCTCTCGCACACCGCCGAGCCAGGATGGAACACCGC 1570

QY 437 ProPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAenIle 456
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1571 ATATTCACCAAGCTGTGTGAGGTACTGGAGAAGCTGCCCAAGTCACTGGGCATTGATGTG 1630

QY 457 GluLeuLysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAenLeuSerThrTyr 476
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1631 GAGATTAAGTGG---CCACAGCGTCCCGAGGCG---GGAGGATCAGAGGCTGAGCAACA 1684

QY 477 PheAspMetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAenSerGlyLys 496
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1685 ATCGACAAGAATCTTTCGCCGAGAGGTGATCATCAGTGATCCAGAAAGGCTGTGGC 1744

QY 497 ArgArgIleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGln 516
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1745 AGACCGAATAATCTTCTCCAGCTTCGATGCTGATGTCACGATGTCAGGTTCAAGCAG 1804

QY 517 AsnLysTyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMet 536
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1805 AAGCTTTCCTGATGATGTTCTCAGCGAGGAGACGAGAGTGGCAGCGCTTCCTTG 1864

QY 537 AspLeuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAenLeuLeu 556
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1865 GATCTCGGAACACGGACCTTTATTCAGCGCGTAACAATGCTCAAGCTTTCGAGCTGGCT 1924

QY 557 GlyIleAsnValHisThrGluAspLeuLeu---ArgAsnProSer---TyrIleGlnGlu 574
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1925 GGCACAGCTCCACACCGCAGGAGCTTCTGGGCGAAAATGCATCAGAAATGCTCGGAAA 1984

QY 575 AlaLysAlaLysGlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAen 594
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 1985 GCAAGGATCTGGCCAAATAGCTGTGATCTGGGCGCAGATTCGAATCCAGGAGCGG 2044

QY 595 ArgArgLysLeuLysGluLeuGlyValAenGlyLeuIleTyrAspArgIleTyrAspTrp 614
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 2045 GTGAGTACTTACCCGGATCGGGCTACCGCCACGCTACGATCGCATGATCTCTTC 2104

QY 615 MetProGlu-----GlnProAenIlePheGlnValGlnLeuGlu---ArgLeuLys 631
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 2105 ATGCGCGGAGGCAAGCGCA-AGCCTTCTTCAATCACCAGCATGATGCGCGGATGCG 2163

QY 632 GlnGluLeuProGluLeuLysSer 639
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 2164 GCGCCAGTCCGGATCTAAGAAAT 2187
```

RESULT 6

```
US-11-097-143-26897
; Sequence 26897, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
```

```
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26897
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-26897
```

Alignment Scores:

Pred. No.:	5,87e-51	Length:	2457
Score:	667.00	Matches:	201
Percent Similarity:	47.24%	Conservative:	124
Best Local Similarity:	29.22%	Mismatches:	230
Query Match:	18.86%	Indels:	134
DB:	12	Gaps:	26

US-10-047-855-3 (1-672) x US-11-097-143-26897 (1-2457)

```
QY 3 ProSerGlnValAlaPheGluIleArg-----GlyThrLeuLeuProGlyGluValPhe 20
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 436 CCACCTCTCGGGAGTCAATGTGCGGCTGGAGGTTCCTCGCCCGCAGGAGCGACTG 495

QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 496 GGGCTTACCGGGGATGTGAAGCCCTCGCGAGTGGCAGCTGTCCAGAGAGTGGCTCTA 555

QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 556 -----GAATCGCTGCAGCAGCTCAACTGGCAGCCACGCTGGCCCTTCAGTCC 603

QY 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 604 TCCCGCCAGCTGGAGTACCGC-----TACTTTGTCTACGTGGAGGATCTCTCT 651

QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 652 GGCTACACAGCAG-----ATCCCGCTGGGAAACCCATTTCAGGCCAGGTCCTCGGA 705

QY 101 Pro----- 101
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 706 CCCTGTACGGAGTACAGTCAGCCAGCTTGGAGCTCTTCGGTATTACGTGGGACAACTCC 765

QY 102 -----LeuGluSerGluIleIleIleLeuAsp----- 109
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 766 GATCTAAAGCCGAGGTGACCGCGCTGGCTGGAACCATGAGGCCCATTCGACGCTCAAG 825

QY 110 -----AspGlyGlnPheGlyIleHisAsnGlyValGluThrLeuAspSerGlyTrp 126
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 826 TTCATATGGGAGAAGATGTTCCAGTCCACAGC-----ATCAGACCTTTGAC----- 873

QY 127 LeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerGluLysProPro----- 144
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 874 -----CCCCAGCACGTCCAGCTGAAGATTGTGCCCGTGGAGAAGCCGCTGGCCTG 924

QY 145 ---ValSerIleThrLysLysLeuLysSerArgPheArgValLysLeuThrLeu 163
|||||:|:| |||||:|:| |||||:|:| |||||:|:|
Db 925 CATGTGGAGTACTCCAAACAGGAGTACGGCAAGAGCCAGCTGGAGCTGAGCTCTTTC 984
```


Qy	3	ProSerGlnValAlaIapheGluIleArg-----GlyThrLeuLeuProGlyGluValPhe	20
Db	4141	CCCACTCTGGCGGAGTTCAATGTGCGGTGAGGTTCCCTTGGCCGCGGAGGAGCGACTG	4200
Qy	21	AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu	40
Db	4201	GGGCTTACCGCGCATGTAAAGCCCTCGGCGAGTGCGAGTGTCCAGAAGTGGCTCTA	4260
Qy	41	LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg	60
Db	4261	-----GAATCGCTGGACGAGCTCAACTGGCAGGCCAGGTGGCCCTTCAGTCC	4308
Qy	61	GlyValSerValGlnTrpArgTyrrPheLysGlyTyrPheLeuGluProLysThrIleGly	80
Db	4309	TGCGCGCAGCTGGAGTACCGC-----TACTTTGTCTACGTGGAGGATCTCTCT	4356
Qy	81	GlyProCysGlnValIleValHilValTrpGluThrHisLeuGlnProArgSerIleThr	100
Db	4357	GGCTACAGCAG-----ATCCGCCGTTGGGAACCCATTTCAGGCCACGTCCTCGGA	4410
Qy	101	Pro-----	101
Db	4411	CCCTGTACGGAGCTACAGTCAGCCAGTGTGACGCTCTTCGGTATTACGTCGACAACTCC	4470
Qy	102	-----LeuGluSerGluIleIleLeuAsp-----	109
Db	4471	GATCTAAAGCCGCGAGTGCCACCGCGCTGGCTGAACCATGAGGCCATTCTGCAGCTCAAG	4530
Qy	110	-----AspGlyGlnPheGlyIleHisAsnGlyValGluThrLeuAspSerGlyTrp	126
Db	4531	TTCAATGGCAGAGATGTTCCAAGTCCACGAC-----ATCGAGACCTTTGAC-----	4578
Qy	127	LeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerGluLysProPro-----	144
Db	4579	-----CCCCAGCAGCTCCAGCTCAAGATTGTGCCGTGGAGAGACCGCTGGCCGTG	4629
Qy	145	---ValSerIleThrLysLysLeuLysLysSerArgPheArgValLysLeuThrLeu	163
Db	4630	CATGTGGAGTACTCCAAACAGGAGTACGGCAAGCAGCAGCTGGAGCTGCAGCCTACTTTC	4689
Qy	164	GluGlyLeuGluAspAspAspAspValSerProThrValLeuHisLysMetSer	183
Db	4690	---GGAGTGCCCTACACCAAGGGGGAATGTCTCATCTTCACATCAGCTGCCGCTGGAG	4746
Qy	184	AsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHisSerGlnPro	203
Db	4747	AGGATGATGGAG-----CAGCAGCTCCGCTTG	4773
Qy	204	GluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGlnThrMetGlu	223
Db	4774	GAGTGC-----TACAGCATGTCCAAC-----	4794
Qy	224	ProAspAsnLeuGluLeuIlePheAspPhePheGluGluAspLeuSerGluHisValVal	243
Db	4795	-----GAACTCTCTG-----	4803
Qy	244	GlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAla	263
Db	4804	-----GGCAGTGCCACTCTGTGTCACTCCGACCTCGGACCTGACT	4836
Qy	264	GluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSer---ArgAsnSerArg	282
Db	4837	-----GGCAGCGAGGGAGTGCTCACTGCGCGATTAGTCGGCCAGATGTCAGAT	4887
Qy	283	LysThrIleGlyLysValArgValAspTyrrIleIleIleLysProLeuProGlyTyrSer	302
Db	4888	GAGACCTTGGCCAGGCTGAGCTTCCCTATGTGCGGTGCGAGCCTTACCGCTACTCGCG	4947
Qy	303	CysAspMetLysSerSerPheSerLysTyrrTrpLysProArgIlePro---LeuAspVal	321
Db	4948	CTAGACTTCAAGAACACCTATGCTCACTACTGGCCCAAGAGCTGGCCCACTGGATGTG	5007

Qy	322	GlyHisArgGlyAlaGlyAenSerThrThrAlaGlnLeuAlaLysValGlnGluAen	341
Db	5008	GGTCATCGTGGAAATGGCAAGAGT---TACATTGCAGACGCTCTCGGAAAGCGAGAAC	5064
Qy	342	ThrIleAlaSerLeuArgAenAlaAlaSerHisGlyAlaAlaPheValGluPheAspVal	361
Db	5065	ACTATCGCGTCTTCTGAGCGGCCATGAACATCAGCAGACATGATCGAGTTGGATGTC	5124
Qy	362	HisLeuSerLysAspPheValProValProValTyrHisAspLeuThrCysCysLeuThrMet	381
Db	5125	CAVTTGACTGCTGATGCGTGTGCTGTGATTTATACGATTTTCGAGCTGCGAACTGCTCCG	5184
Qy	382	LysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThr	401
Db	5185	CCTGGCAAGCAGATACCGCCCCAGACCCAGCTGGAGTACGTGTGTATCAAAAGCATAAAC	5244
Qy	402	PheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArg	421
Db	5245	TATGAGCTGCTCAAAAGGCTGCGCATCTTCTCTGTGATCGCA-----	5286
Qy	422	LysGluSerValValGlnGlnGluAenSerPheSer-----GluAenGln	436
Db	5287	-----GGCCAAAGTGAGGGAGTATCCTCTCGCAACAAGCCGAGCCAGGATGGAAACACCGC	5340
Qy	437	ProPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAenIle	456
Db	5341	ATATTTCCAAACGCTGTGTGGAGGTACTGGAGAACTGCCCCAAGTCACTGGCGCATGTGATG	5400
Qy	457	GluIleLysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAenLeuSerThrTyr	476
Db	5401	GAGATTAAAGTGG--CCACAGCGTCGCCAGGCG--GGAGGATCAGAGGCTGAGCAACA	5454
Qy	477	PheAspMetAenLeuPheLeuAspIleIleLeuLysThrValLeuGluAenSerGlyLys	496
Db	5455	ATCCACAAGAAACTTCTTCGCCGCAAGAGGTATCCATCAGGTGATCCAGAAGGCGTGTGGC	5514
Qy	497	ArgArgIleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGln	516
Db	5515	AGACCGATAATCTTCCACGCTTCGATGCTGACATGTGCACATGCTGAGGTTCAACGAG	5574
Qy	517	AsnLysTyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMet	536
Db	5575	AACGTCTTCCAGTGTATGTTCTTCACGCGGAGAGAGCAAGAAGTGGCAGCGTGTCTCTG	5634
Qy	537	AspLeuArgSerArgThrThrProIleAlaMetSerPheIleGlnPheGluAenLeuLeu	556
Db	5635	GATCTGCGAACAACGACCTTTATTTGCAAGCCGTAACAATGCTCAAGCTTTCGAGCTGGCT	5694
Qy	557	GlyIleAenValHisThrGluAspLeuLeu---ArgAsnProSer---TyrIleGlnGlu	574
Db	5695	GGCACAGCTCCACAGCCGAGAGACTTCTCTGGGCGGAAAATGCATCAGAAATGCTCGCGNAA	5754
Qy	575	AlaLysAlaLysGlyLeuValIlePheCysTrpGlyAspAspThrAenAspProGluAen	594
Db	5755	GCCAAGGATCTGGGCGAAATAGCTGTGATCTGGGGCGGACGATTGCAACTCCAAGGAGCGG	5814
Qy	595	ArgArgLysLeuLysGluLeuGlyValAenGlyLeuIleTyrAspArgIleTyrAspTrp	614
Db	5815	GTGCGACTTTCACCCGGATTCGGGGCTTACCGCCACGCTGCTACCATGCGACTGATCTCTTC	5874
Qy	615	MetProGlu-----GlnProAenIlePheGlnValGlnLeuGln---ArgLeuLys	631
Db	5875	ATCCGCGAGGGCAAGCGCGA-AGCCTTCTTCAATCAACCCGCACTGATGCCCGGATTCGC	5933
Qy	632	GlnGluLeuProGluLeuLysSer	639
Db	5934	GGCCCAAGTGGCGGATCTAAGAAAT	5957

RESULT 8

US-11-097-143-27181
; Sequence 27181, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig

```

; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 27181
; LENGTH: 9667
; TYPE: DNA
; ORGANISM: DROSOPHILA
; US-11-097-143-27181

```

Alignment Scores:		
Pred. No.:	3,78e-50	Length: 9667
Score:	667.00	Matches: 201
Percent Similarity:	47.24%	Conservative: 124
Best Local Similarity:	29.22%	Mismatches: 230
Query Match:	18.86%	Indels: 134
DB:	12	Gaps: 26
NC-10-047-985-3 (1-672) x	US-11-097-143-27181 (1-9667)	

QY	3	ProSerGlnValAlaPheGluIleArg-----GlyThrLeuLeuProGlyGluValPhe	20
DB	6646	CCCACTCTCGGGAGTTCAATGTGGCTGGAGGTTCCCTTGGCCGCCGAGGACGACTG	6705
QY	21	AlaIleCysGlySerCysAspAlaIleGlyAsnTrpAsnProGlnAsnAlaValAlaLeu	40
DB	6706	GGGCTTACCGCGATGTAAAGCCCTCGCGAGTGGCAGCTGTCACGAAGTGGCTCTA	6765
QY	41	LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg	60
DB	6766	-----GAATCGCTGGACGAGCTCAACTGGCAGGCCACGGTGGCCCTTCAGTCC	6813
QY	61	GlyValSerValGlnTrpArgTrpPheLysGlyTrpPheLeuGluProLysThrIleGly	80
DB	6814	TGCCGCCACGCTGGAGTACCGC-----TACTTTGTCTACGTGGAGATCTCTCT	6861
QY	81	GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr	100
DB	6862	GGCTACAGCAG-----ATCCGCGGTTGGGAAACCCATTTCAGGCCAGGTCCCTGGGA	6915
QY	101	Pro-----101	
DB	6916	CCCTGTACGGAGCTACAGTGCACCGCAGTTGGACGCTTTCGGTATTACGTGGACACTCC	6975
QY	102	-----LeuGluSerGluIleIleAsp-----109	
DB	6976	GATCTAAAGCCGAGGTGCACCGCGCTGGCTGAACCATGAGGCCATTCTGCAGCTCAAG	7035
QY	110	-----AspGlyClnPheGlyIleHisGlnGlyValGluThrLeuAspSerGlyTrp	126
DB	7036	TTCAATGGCGAGAAGATGTTCCAAGTCCACGAC---ATCGAGACCTTTTGAC-----	7083

127	LeuThrCysGlnThrGluLeuArgLeuGluHisTyrSerGluLysPro	144
7084	-----CCCAGCACGTCGACGTGAAGATTGTCCTGGAGAACCGCTGCCTG 7134	
145	---ValSerIleThrLysLysLysLysLysSerArgPheArgValLysLeuThrLeu 163	
7135	CATGTGGAGTACTCCAAACAGGAGTACGGCAAGAGCCAGCTGGAGCTGACGCTACTTTC 7194	
164	GluGlyLeuGluAspAspAspAspValSerProThrValLeuHisLysMetSer 183	
7195	---GGAGTGCCTACACCAAGGGGACATTGTCTTTCATCATCACGCTGCCGTGGAG 7251	
184	AsnSerLeuGluLeuSerLeuIleSerAspAsnGluPheLysCysArgHisSerGlnPro 203	
7252	AGGATGATGGAG-----CAGCACTTCGCTGTG 7278	
204	GluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGlnThrMetGlu 223	
7279	GAGTGC-----TACAGCATGTCCAAC----- 7299	
224	ProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGluHisValVal 243	
7300	-----GAATCTCTG----- 7308	
244	GlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAla 263	
7309	-----GGCAGTGCACCTCTGTCTACCTCGACCTGACT 7341	
264	GluSerGlyLysSerAlaGlyLeuLeuThrLeuProIleMetSer---ArgAsnSerArg 282	
7342	---GGCACGGAGGAGTGCTCCACTGCCGATTAAAGTCGGCAAGAATGCAGAT 7392	
283	LysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGlyTyrSer 302	
7393	GAGACCTGGCCAGGCTGAGGCTTCCCTATGTCGGTGCAGCTTACCCTACTCGCCG 7452	
303	CysAspMetLysSerSerPheSerLysTyrTrpLysProArgIlePro---LeuAspVal 321	
7453	CTAGACTTCAAGAAACACCTATGTCTACTTCTGGCCCAAGAGCTGGCCCAACCTGGATGTG 7512	
322	GlyHisArgGlyAlaGlyAsnSerThrThrThrAlaGlnLeuAlaLysValGlnGluAsn 341	
7513	GGTCATGCTGGAATGGCAAGAT---TACATTGCAGACGCTCTCTGCGGAAGGAGGAAC 7569	
342	ThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAspVal 361	
7570	ACTATCGGCTCTTCTCTGAGCGCCCATGAACATCACGACACATGATCGAGTTGGATGTC 7629	
362	HisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMet 381	
7630	CATTTGACTGTGTGTGTGCTGTGATTTATCAGGATTCGAGCTCGCAGTCTGCTCCG 7689	
382	LysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThr 401	
7690	CCTGGCAAGCAGATCAGCGCCCGACACCGCTGGAGTACGTGCTGATCAAGACATAAAC 7749	
402	PheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArg 421	
7750	TATGAGCTGCTCAAAAGGTGCGCATCTCTCTGTGTGTCGA----- 7791	
422	LysGluSerValValGlnGluGluAsnSerPheSer-----GluAsnGln 436	
7792	-----GGCCAAGTGAGGAGTATCCCTCGCACACCGCGAGCCCGAGTGAACACCGC 7845	
437	ProPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIle 456	
7846	ATATTCCCAACGCTGTGTGAGGTACTGGAGAAGCTGCCCAAGTCACTGGCGATTGATGTG 7905	
457	GluIleLysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyr 476	
7906	GAGATTAAAGTGG---CCACAGCGTCCCGAGG---GGAGGATCAGAGCTGAGCAACA 7955	
477	PheAspMetAsnLeuPheLeuAspIleLeuLysThrValLeuGluAsnSerGlyLys 496	

```
Db 7960 ATGCACAAAGTCTTTCGCCACCAAGGTGATCCATCATGATCCAGAAAGGCGTGTGC 8019
Qy 497 ArgArgIleValPheSerPheAspAlaAspIleCysThrMetValArgGlnLysGln 516
Db 8020 AGACCGAATAATCTTCCAGCTTCGATGCTGACATGTGCACATGCTGAGGTTCAGCAG 8079
Qy 517 AsnLysTyrProIleLeuPheLeuThrGlnGlnLysSerGluIleTyrProGluLeuMet 536
Db 8080 AACGTCCTCCAGTAGTATCTCCACGAGGAGAGACGAAGAGTGGCAGCGCTTCCTG 8139
Qy 537 AspLeuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeu 556
Db 8140 GATCTGCGAACAACGACCTTTATGACGCGGTAAACAATGCTCAAGCTTTCGAGCTGCT 8199
Qy 557 GlyIleAsnValHieThrGluAspLeuLeu---ArgAsnProSer---TyrIleGlnGlu 574
Db 8200 GGCACAGCTCCACAGCCGAGGACTCTCTGGCGGAATGCAATGATCGAGAA 8259
Qy 575 AlaLysAlaLysGlyLeuValIlePheCysThrGlyAspAspThrAsnAspProGluAsn 594
Db 8260 GCCAAGGATCTGGGCAATAGCTGTGATCTGGGGCGACGATTGCAACTCCAAAGGAGCG 8319
Qy 595 ArgArgLysLeuLysGluLeuGlyValaAngLysLeuIleTyrAspArgIleTyrAspTrp 614
Db 8320 GTGCAGTACTTCAACCGGATCGGGCTACCGCCACGTGCTACGATCGCAGTGATCTCTTC 8379
Qy 615 MetProGlu-----GlnProAsnIlePheGlnValGlnGlnGlu---ArgLeuLys 631
Db 8380 ATGCCGAGGGCAAGCGCA-AGCTTCTTCAATCACCACCGCAGTGTGGCGGAGTTGCG 8438
Qy 632 GlnGluLeuProGluLeuLysSer 639
Db 8439 GSCCCAGTCCGGATCTAAGAAAT 8462
```

RESULT 9

```
US-11-097-143-19967
; Sequence 19967, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19967
; LENGTH: 2299
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-19967
Alignment Scores:
Pred. No.: 6,646-51 Length: 2299
```

```
Score: 666.00 Matches: 177
Percent Similarity: 50.00% Conservative: 133
Best Local Similarity: 28.55% Mismatches: 250
Query Match: 18.83% Indels: 60
DB: 12 Gaps: 17
US-10-047-855-3 (1-672) x US-11-097-143-19967 (1-2299)
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTyrAsnProGlnAsnAlaValAlaLeu 40
Db 416 GCCCTGTTGGGCAATCTCCAGTGTGGCGCTGCAAGCGGACGGCTGTGTCTCTC 475
Qy 41 LeuProGluAsnAspThrGlyGluSerMetLeuTyrPheAlaThrIleValLeuSerArg 60
Db 476 -----AACCGCACTCAATCTGAATCTTGGAGCGCTCGTGGAGATACCGCAA 526
Qy 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 527 AACAGCAGCTGTGATGCTACTTTGCCGCC-----GCCGTTGGC 568
Qy 81 GlyProCysGlnValIleValHieThrLysTyrGlnThrHisLeuGlnProArgSerIleThr 100
Db 569 CAGAGTGGAGCGCTGCAGATTCCGCGCTGGGAGTGCATGTCCAGGCAAGGACCTTAAAT 628
Qy 101 ProLeuGluSerGluIleIleLeaAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
Db 629 ACCACCAAGTTCACTGGCAATCGCAGCGAT---GAGTTCCGACTCATTTGGCAATGAGCGC 685
Qy 121 ThrLeuAspSerGlyTyrLeuThrCysGlnThrGluIle-----Arg 134
Db 686 CAGTTGTCCGCTGGCTGCTCAGGACTCGGGCAACATTTGCACTTGAAGCTCTTCCGC 745
Qy 135 LeuArgLeuHisTyrSerGluLysProValSerIleThrLysLysLysLysLys 154
Db 746 GAGGCTCTGAGCTTGGAGAGCAACCGAGGTGGGCGAGAGAGCTGCTGCGT--- 802
Qy 155 SerArgPheArgValLysLeuGluGlyLeuGluGluAspAspAspArgVal 174
Db 803 -----TTGCAACCACTAGATCCAAAGACTCTTA 829
Qy 175 SerProThrValLeuHieLysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsn 194
Db 830 GCGCC---ATCCTAAGTTCCGACGTGCCATGTGGAGTACGTGCGAATGGCGTACG 886
Qy 195 GluPheLysCysArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrp 214
Db 887 GACAGTAAGTCCGG---GTGCAGCCGAGTACGAGTGCCCTACGCCACCGGTGACATC 943
Qy 215 ThrGluTyrSerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPhePhe 234
Db 944 CTCATGTTCCATGTGAACCTGGAACAGCTGGATCGAGTGGCGCTATCTGTGGAAGTTAT 1003
Qy 235 GluGluAspLeuSerGluHieValValGlnGlyAspAlaLeuProGlyHisValGlyThr 254
Db 1004 GCCGAACCAACCGGATGAGGACCTGCTCGCT-----CTCTAGGATAC 1045
Qy 255 AlaCysLeuLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeu 274
Db 1046 TCGCACCTGTGCACACCTCGTTC-----TCCGCGC---ACTGAAGGCGATGTGTCATC 1096
Qy 275 ProfileMetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIle 294
Db 1097 GACCTTATTTCCGCGCTCTGGCAGCGCTGGTGGGTCAACTCGCTATTGAGTACCTGCAC 1156
Qy 295 IleLysProLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLys 314
Db 1157 ATTGCGCCCATGGAGAGGATGGCTTCGATCTGCGCACACGCTTCGCCAACACTACTGGCGC 1216
Qy 315 ProArg---IleProLeuAspValGlyHisArgGlyAlaGlyAsnSer---ThrThrThr 332
Db 1217 TCCAAATGGACAGCCCTGGAGATTGGCCATCGGGATTGGGAAATCTCTTACGTCACC 1276
Qy 333 AlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHis 352
```



```

; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19966
; LENGTH: 4299
; TYPE: DNA
; ORGANISM: DROSOPHILA
; US-11-097-143-19966

Alignment Scores:
Pred. No.: 1,56e-50      Length: 4299
Score: 666.00           Matches: 177
Percent Similarity: 50.00%      Conservative: 133
Best Local Similarity: 28.55%    Mismatches: 250
Query Match: 18.83%             Indels: 60
DB: 12                         Gaps: 17

US-10-047-855-3 (1-672) x US-11-097-143-19966 (1-4299)

Qy 21 AlaileCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValaLeu 40
Db 1416 GCCTGTGTGGCAATCTGCAGTCTGGCGCTTGGCAACGGGAGCGCTGTGCTCTC 1475
Qy 41 LeuProGlnAsnAspThrGlyGluSerMetLeuTrpLysAlaThrLeuValLeuSerArg 60
Db 1476 -----AACCCGACTGCATCTCAATGTTTGGAGCGCTCCGTGGAGATACCGCAA 1526
Qy 61 GlyValSerValGlnValIleValHisLysTrpGlnThrHisLeuGlnProArgSerIleThr 80
Db 1527 AACAGCAGCGTTGAGTATCGCTACTTTGCCGCC-----GCCGTTGCC 1568
Qy 81 GlyProCysGlnValIleValHisLysTrpGlnThrHisLeuGlnProArgSerIleThr 100
Db 1569 CAGAGTGGAGCCGTGCAGATTCCGCCGTGGAGTCCGATGTCAGGCAAGACCTTTAAT 1628
Qy 101 ProLeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
Db 1629 ACCACCAAGTTCAGTGGCAATCGCAGCGAT---GAGTTCGGACTCATTTGGCAATGAGCGC 1685
Qy 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIle-----Arg 134
Db 1686 CAGTTGTCCGCTGGCTGGCTTTCAGGACTCGGGCAACATTTGCGAGTTGAAGACTTTCGCG 1745
Qy 135 LeuArgLeuHisLysTrpSerGluLysProProValSerIleThrLysLysLysLysLys 154
Db 1746 GAGGCTCTGAGCTGGAGGACCAACCGAGGTGGGGCAGGAGAGAGCTCGCTCGCT--- 1802
Qy 155 SerArgPheArgValLysLeuThrLeuGluGlyLeuGluGluAspAspAspArgVal 174
Db 1803 -----TTGCAACCAAGTAGATCCAAAGACTCTA 1829
Qy 175 SerProThrValLeuHisLysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsn 194
Db 1830 GGGCCC---ATCCTAAGTTCCGACAGTGCCTAATGTGGAGTAGCTGCGAATGCGGTACCGT 1886
Qy 195 GluPheLysCysArgHisSerGlnProGluCysGlyTrpGlyLeuGlnProAspArgTrp 214
Db 1887 GACAGTAAGCTCGG---GTGACGCGCGGAGTACGAGTGCCCTACGCCACCGGTGCATC 1943
Qy 215 ThrGluTrpSerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPhePhe 234

```

```

RESULT 10
US-11-097-143-19966
; Sequence 19966, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05

```


355	Db	-----GATGAGTGTATATACGCTTCTGGGAGTCCCAATATATATCCAGAGTG	1378	ATGGGAGGATCAAAACGAAAGACCTCATTCGCTCCCTGGAGGCAATTTTCTTCAGAT	1437
99	Qy	IleThrProLeuGluSerGluIleIleAspGlyGlnPheGlyIleHisAsnGly	404	GlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGlu	423
403	Db	ATTCGAACCTTGTCAGAACATCTCTAGAACTGCGACGCTTTGGAAAACACACAGCAT	1438	CAACTGAAGAGGTGAAGTCTCGCATCTCGCTGCACCAAGGCTGCGATAAG	1491
119	Qy	ValGluThr-----LeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArg	424	SerValValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMet	443
463	Db	GATGAAGCGAACACAGGTGATCGCGCTGGCCACACAGAGACAATTTGTCACCTGAAG	1492	-----TCTGTGATCGAATGCTGTGGAGCAGCGCCCTTTCCCTGCTCTCCGAC	1542
137	Qy	LeuHisTyrSerGluLysProProValSerIleThrLysLysLysLysLysSerArg	444	ValLeu-----GluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIle	461
523	Db	ATCTTC-----AATGCTCCATCTCTGTGGCAACGTCAGAACCCGAGCTCTCTAT	1543	CTCTCGATGAGGAAATATACCGCTCGACATGGGCTTCTTAATCGAGATCAAGTGGCG	1602
157	Qy	PheArgValLysLeuThrLeuGluGluLeuGluAspAspArgValSerPro	462	CysGlnGlnArgAspGlyMetTrpAsp---GlyAsnLeuSerThrTyrPheAspMetAsn	480
574	Db	GTCCATGTGACCCCATGTTGAGTGGCGGAGAACTCTTGAAGAACACAGCAATCCC	1603	CAGATGACCAATATGCGCGTGGGAGAGCGGAGCTTTAAGCCACGCTTCGATCGAAG	1662
177	Qy	ThrValLeuHisLysMetSerAsnSerLeuGluIleSer---LeuIleSerAspAsnGlu	481	LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal	500
634	Db	-----ATCAAAATGTTTCGTCGCAACACGCTTTATCGCTTATCTCAGCACCGGAA	1663	TTCTATGTGGACACCATATTTGGAGATTGTTTAAATAAGGCTGGAAGAGGCGCATAGTC	1722
196	Qy	PheLysCysArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThr	501	PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro	520
688	Db	ATTAGGCGAGGAACAGCATCTTCAA-----TTGGCAAAAGTCGAGGTGACC	1723	TTTTCGAGCTTCGATCCGACATCTGTGCATGTTTGGTTTAAAGCAGAAATGTATATCCC	1782
216	Qy	GluTyrSerIleGlnThr-----MetGluPro	521	Ile---LeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArg	539
736	Db	AATCTATGTGTCAAAATCTTTGGCGCGGACGACAGATTGGAGCCAGATGGGCGG	1783	GTGACTTTACTCTGAGGATCCGACTCCCGGTT-----CAGTATGCAGACCAAGG	1836
225	Qy	AspAsnLeuGluLeuIle-----PheAspPheGluGluAspLeuSerGlu---	540	SerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuLysLeuAsn	559
796	Db	AAGGATATGAGCTATTTTCAATGTTCCATCGCTTTTCGGGAGGAAACGCTCTACCGACTG	1837	GTAAGCTGCGAGATGTTGCGGAGGTTTGTCAACAGCTGTGAGTTCTTGGGCTGACC	1896
241	Qy	-----HisValValGlnGlyAspAlaLeuProGlyHisValGlyThr	560	ValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGly	579
856	Db	GATTTGTACATACGCGCATTAAGCAGCTATGATGAACACCGTATCATTCACGGTAC	1897	CTTCATCTCAACTCACTGCTGAAACAAGCCCTCGACTATGGCATCTGTCACCATCAAT	1956
255	Qy	AlaCysLeuLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeu	580	LeuValIlePheCysTyrTrpGlyAspThrAsnAspProGluAsnArgArgLysLeuLys	599
916	Db	GGATTCCTGTCACGATCAGCTGGCGGACCGAAGGCTCTGCCCGGTG-----	1957	CTGGACGCTTTGTTGTTATGTTGTTCCACCATTCGCTGAGGATCCGCAATAGCTAAAG	2016
275	Qy	ProIleMetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIle	600	GluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnPro	619
967	Db	AGATCACTTGGCGCTCCACCCGCTCCACTTAATGAGATGTGTACGCTATCTAATA	2017	AAGCAGGAGTGTGGGAAATACTACGATCGTCTC-----	2052
295	Qy	IleLysProLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrp---	620	AsnIlePheGlnValGlnLeuGluArgLysGlnGluLeuProGluLeuLysSer	639
1027	Db	ATTCGACCTTGGCAAAATTTCCGTTGCGATTGAGCCACAGCTACGAGCGTTACTGGCGC	2053	-----GACCAACTGGACCAAGTGGGCGAAGAGCTGGAGGGGACACCATG	2097
314	Qy	LysProArgIleProLeuAspValGlyHisArgGlyValGlyAsnSerThrThrThrAla	640	CysLeuCysProThrValSer	646
1087	Db	AAAAATCGCTATGCTGAACATCGGGCAACAAAGGATCGGGAAT-----ACGTAC	2098	TGCACCATTTGATTGAGTGACC	2118
334	Qy	GlnLeuAlaLys-----ValGlnGluAsnThrIleAlaSerLeuArgAsnAlaLysSer	RESULT 12		
1138	Db	CGGTTAGATCCGATGTTGAGGAGAAATACCTGTATGTTTCAAGCAGGCTGTTTG	US-11-097-143-40594		
352	Qy	HisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheValProValVal	Sequence 40594, Application US/11097143		
1198	Db	GCCAAATGACAGATGGTGGAGATGATGTCAGCTCAGCAGGATGCCAGGTTGGTGGTG	GENERAL INFORMATION:		
372	Qy	TyrHisAspLeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspProVal---	APPLICANT: Venter, J. Craig		
1258	Db	TATCAGATTTTGTACTGCGATTATGTTGAGAGAAATGCTAGTTTGGAGTCTCTCTG	TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID		
390	Qy	-----	TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE		
1318	Db	GAGAAATCAGGATCTGCTGATATTTCCTACGAAATCTTAACAAATCAATGCTCCTCGCT	TITLE OF INVENTION: DROSOPHILA GENES.		
391	Qy	-----GluLeuPheGluLeuProValLysGluLeuThrPheAsp	FILE REFERENCE: CLO00728		
		-----	CURRENT APPLICATION NUMBER: US/11/097,143		
		-----	PRIOR FILING DATE: 2005-04-04		
		-----	PRIOR APPLICATION NUMBER: 60/157,832		
		-----	PRIOR FILING DATE: 1999-10-05		
		-----	PRIOR APPLICATION NUMBER: 60/160,191		
		-----	PRIOR FILING DATE: 1999-10-19		
		-----	PRIOR APPLICATION NUMBER: 60/161,932		
		-----	PRIOR FILING DATE: 1999-10-28		
		-----	PRIOR APPLICATION NUMBER: 60/164,769		

Qy 540 SerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAenLeuLeuGlyIleAen 559
Db 2837 GTAACGTCGAGGATGTTGCGTGGAGTTTGGCAACAGTCTGGAGTTCTTGGGCTGACC 2896
Qy 560 ValHisThrGluAenLeuArgAenProSerTyrlleGlnGluAlaLysAlaLysGly 579
Db 2897 CTTTCATGCTAATCACTGCTGCAACAGCCCTCGACTATGGCATACCTGGCACCAGATCAAT 2956
Qy 580 LeuValIlePheCysTrpGlyAspThrAsnAspProGluAenArgLysLeuLys 599
Db 2957 CTGGACGCTTTGTTATGGATGTTCCACCATTCACCTGGAGATCCGCAATAAGCTAAAG 3016
Qy 600 GluLeuGlyValAsnGlyLeuIleTyrlleTyrlleArgIleTyrlleAspTrpMetProGluGlnPro 619
Db 3017 AGCAGCGAGTCTGGGATATATACGATGCTCTC----- 3052
Qy 620 AsnIlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSer 639
Db 3053 -----GACCAACTGGACCAAGTGGCGGAAGAGCTGGAGGGGGACACCATG 3097
Qy 640 CysLeuCysProThrValSer 646
Db 3098 TGCACATGATTCACTGACC 3118
RESULT 13
PCT-US05-00517-1279
; Sequence 1279, Application PC/TUS0500517
; GENERAL INFORMATION:
; APPLICANT: THE OHIO STATE UNIVERSITY
; TITLE OF INVENTION: METHODS OF USING DATABASES TO CREATE GENE-EXPRESSION MICROARRAYS,
; TITLE OF INVENTION: MICROARRAYS CREATED THEREBY, AND USES OF THE MICROARRAYS
; FILE REFERENCE: 18525-04130
; CURRENT APPLICATION NUMBER: PCT/US05/00517
; CURRENT FILING DATE: 2005-01-07
; PRIOR APPLICATION NUMBER: 60/535,111
; PRIOR FILING DATE: 2004-01-08
; NUMBER OF SEQ ID NOS: 3859
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1279
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Equus caballus
PCT-US05-00517-1279
Alignment Scores:
Pred. No.: 1.08e-48 Length: 693
Score: 634.50 Matches: 120
Percent Similarity: 96.06% Conservative: 2
Best Local Similarity: 94.49% Mismatches: 4
Query Match: 17.94% Indels: 1
DB: 1 Gaps: 1
US-10-047-855-3 (1-672) x PCT-US05-00517-1279 (1-693)
Qy 547 MetSerPheAlaGlnPheGluAenLeuGlyIleAenValHisThrGluAenLeuLeu 566
Db 2 ATGAGCTTTGCACAGTTTGAATCTACTTGGGATTAATGCACATACCTGAAGACCTGCTC 61
Qy 567 ArgAsnProSerTyrlleGlnGluAlaLysGlyLeuValIlePheCysTrpGly 586
Db 62 AGAAACCCGCTCTCATTCACAGAGGCAAGCTAAGGGACTAGTGCATATTTCTGCTGGGT 121
Qy 587 AspAspThrAsnAspProGluAenArgLysLeuLysGluLeuGlyValAsnGlyLeu 606
Db 122 GATGATACCAACGACCCCTGAAACAGAAAGAAATTTGAAGGAATTTGGAGTTAATGCTCTA 181
Qy 607 IleTyrlleArgIleTyrlleAspTrpMetProGluGlnProAsnIlePheGlnValGluGln 626
Db 182 ATTTAATGATGATATATGATTGGTTCCTGACAGCCCAATATATATTCAGGTGAGCAG 241
Qy 627 LeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCysLeuCysProThrValSer 646

Db 242 TTGGAACGCTGAACAAGAAATTGCCAGAGCTTAAGAGCTGTTGTGTCCCACTGTTAGC 301
Qy 647 ArgPheValProSerSerLeuCys---GlyGluSerAspIleHisValAspAlaAsnGly 665
Db 302 CGCTTTGTTCCCTCATCTTTGTGTGGTGGGAGCCTGACATCCATGTGATGCCAATGCG 361
Qy 666 IleAspAsnValGluAsnAla 672
Db 362 ATCGATAACGTGGGAGTGTCT 382
RESULT 14
US-11-097-143-33686
; Sequence 33686, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33686
; LENGTH: 2100
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-33686
Alignment Scores:
Pred. No.: 3.28e-40 Length: 2100
Score: 550.00 Matches: 182
Percent Similarity: 44.85% Conservative: 110
Best Local Similarity: 27.96% Mismatches: 261
Query Match: 15.55% Indels: 100
DB: 12 Gaps: 21
US-10-047-855-3 (1-672) x US-11-097-143-33686 (1-2100)
Qy 14 LeuLeuProGlyGluValPheAlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsn 33
Db 256 CTGGCCCGCCGATGAGTTGTGGCGGATCATCCGCTCTGGGCTGTGGCGCAG 315
Qy 34 ProGlnAsnAlaValAlaLeuLeuProGluAenAspThrGlyGluSerMetLeuTrpLys 53
Db 316 CTCGAGAGGCTCTACCCCTCAAGGAGCAGGACGAGCGCGCACCAAGTGGTATCTGAGG 375
Qy 54 AlaThrIleValLeuSerArgGlyValSerValGlnTyrlleArgTyrllePheLysGlyTyrllePhe 73
Db 376 GTTTGGATATGCCGCCAGCCAG-----AGATTCTACTACCGATACCTG 417
Qy 74 LeuGluProTyrlleTyrlleGlyProCysGlnValIleValHisLysTrpGluThrHis 93
Db 418 ATATATCTCCAAAGAACAGCAAGGC-----AATCGGGTCTCTGAGGACCTGGGAGGCGCCAG 471
Qy 94 LeuGlnProArgSerIleThrProLeuGlu-----SerGluIleIle 107

```
Db 472 AGAGTGGCTGAATGTCAGACCTACGAGATCTACCGCTCCCGGGTCTGGACATCTTC 531
Qy 108 IleAspAspGlyGlnPheGlyIleHisAsnGlyValGluThrLeuAspSerGlyTyrLeu 127
Db 532 GCGGAGCCCTACCCGCTCAGTGGCGGGGATTC---TATCTGGNACCGGTTGGTTG 588
Qy 128 ThrCysGlnThrGluLeuArgLeuGluHisTyrSerGluLysProProValSerile 147
Db 589 CAGTACGAGTACGTGATAGAGTAAATTCGTCGCGAGCATATATGCTATTCT--- 645
Qy 148 ThrLysLysLysLeuLysLysSerArgPheArgValLysLeuThrLeuGluLeuGlu 167
Db 646 -----GGAAATGGGAAGTAACGCCAAGTACCGGTTGATCTTACACCGCTCAAT 693
Qy 168 GluAspAspAspArgValSerProThrValLeuHisLysMetSerAsnSerLeuGlu 187
Db 694 TTGATAGACGACACATAT-----GAG 717
Qy 188 IleSerLeuIleSerAspAsnGluPheLysCysArg-His---SerGlnProGluCysG1 206
Db 718 GTCTCCAGATATGCTTACATCATCATCTTCGCGATTTCACCAACCGCTGGACACC-TACAAATGC 777
Qy 206 yTyrGlyLeuGlnPro-----AspArgTyrThrGluTyr----- 217
Db 778 TACACACAAGGACCACTATTGTTGCTTCGCGATTTCACCAACCGCTGGACACC-TACAAATGC 836
Qy 218 -----SerIleGlnThrMetGluProAspAsnLeuLeuIlePheAspPhePh 234
Db 837 CCTTCGCTGCTATTATATCAGGCATCAAGAT-----CTCAGTTA----- 879
Qy 234 eGluGluAspLeuSerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyTh 254
Db 880 -----TCCCTTGGAGAGCGGTACATATTTCCGGATCAATCAAGGCG----- 921
Qy 254 rAlaCysLeuLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLe 274
Db 922 -----AGTCGTGGGATTTTGCAGCT 941
Qy 274 uProIleMetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIle11 294
Db 942 CCCATCTTTGTAGTTTACAAATTTGGCCATTTGGGAGATCAGCTTGCCATTTATTTGGT 1001
Qy 294 eLysLysProLeuProGlyTyrSerCys---AspMetLysSerSerPheSerLysTyrTr 313
Db 1002 GGTTCACCTATGCCAAAAGTGGAGCGGGAATCTACGAGCCAGTTTCCATCACTACTG 1061
Qy 313 pLysProArgIlePro---LeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrTh 332
Db 1062 GCGCGATAATTGGCCACATTTGGATGGGATACCGAGGTTTGGAGCCAGCTTACTTCCA 1121
Qy 332 rAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHi 352
Db 1122 ATCTTCA---ACGAGTCTTACGAGAACACCATCGAAAGATTATCTGGCAGTTCTGAAGGC 1178
Qy 352 sGlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheValProValValTy 372
Db 1179 CAAGGAGATATGTTCCAACTGGATGGCAGTGGACCAAGGATACGTTCCCGTAGTCTG 1238
Qy 372 rHisAspLeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspProValGluLe 392
Db 1239 GCATGATTTGGATTTTACACCTCAGACACAGATCGCTCTGTAAGGATCGCTTTGATTT 1298
Qy 392 uPheGluIleProValLysGluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHi 412
Db 1299 ACGTTTGTGCTAATAAGGAACCTTACTACTCGAA-----CTCAAGCCGACCGC 1349
Qy 412 sValThrAlaLeuLysSerLysAspArgLysGluSerValValGlnGluAsnSerPh 432
Db 1350 GGTGTTATCTTC-----AAACGTTGGACGCTCAGATACACCAATCT 1394
Qy 432 eSer-----GluAsnGlnProPheProSerLeuLysMetValLeuG1 446
Db 446 -----
```

```
Db 1395 TAATGTCAAGGATGTGAGCCAAACACAGAAATTTTCCAAAACTCTCGAGGTTTTTGA 1454
Qy 446 uSerLeuProGluAspValGlyPheAsnIleGluIleLysTyrIleCysGlnGlnArgAs 466
Db 1455 GGCCTTACCACAAACACTGGGACTTCTGTGGAGATTAAGTGG-----CCGCAAT-AA 1507
Qy 466 pGlyMetTyrAspGlyAsnLeuSerThrTyrPheAspMet-AsnLeuPheLeuAspIleI 486
Db 1508 TGGCTTCTGTGTGCGGAATCCACCAGAGTTTGAACAAAAACATTTACGTGGACAGAA 1567
Qy 486 leLeuLysThrValLeuGluAsnSerGlyLysArgIleValPheSerSerPheAspA 506
Db 1568 TTTACAGATCAGATTCATCATGGGTGTGGAGCATCCACTAATCTTCGCAGCTTTTATG 1627
Qy 506 laAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeuThrG 526
Db 1628 CGATATATGCACAAATGATTAGGCTGAGCAGCATGTCTTCCTGTGATTCTGTAGACA 1687
Qy 526 lncGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThrProIleA 546
Db 1688 TTGGCAAGTCTCAGATTTGGGATGAGTATATGATCTGAGGGCAGACAGATTCCAGCAAG 1747
Qy 546 laMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnValHisThrGluAspLeuL 566
Db 1748 CCATCAACTCGTTCAGTCGCGACAGAAATTTTGGAAACAGCTCTGCACGTGGAACACTTC 1807
Qy 566 euArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTyrG 586
Db 1808 AAAACAAACACGACGCAACTTGGCCTTGGATCTGCAGCAGTCTCTTTTTTATGGG 1867
Qy 586 lyAspAspThrAsnAspProGluAsnArgLysLysLysGluLeuGlyValAsnGlyL 606
Db 1868 GTAACGATATGCAAGATGAACATCTTCTGGAGCAATTTAGAGCTCTGGATGTGACGGGTC 1927
Qy 606 euIleTyrAspArgIle-----TyrAspTyrMetProGluGlnProAsnI 621
Db 1928 TGATTTTACCATCATCGATCGAGTCGGTCCGTTCCGTTGG-----AAACGATCCGAAT 1981
Qy 621 lePheGlnValGluGlnLeuGluArgLeu 630
Db 1982 TTTTCCGGGCCCCACAACTTATGGAACATA 2010

RESULT 15
US-11-097-143-33685
; Sequence 33685, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33685
```


Tue Jul 5 14:07:50 2005

! LENGTH: 4100

! TYPE: DNA

! ORGANISM: DROSOPHILA

US-11-097-143-33685

Alignment Scores:

Pred. No.: 8,146-40 Length: 4100
Score: 550.00 Matches: 182
Percent Similarity: 44.85% Conservative: 110
Best Local Similarity: 27.96% Mismatches: 261
Query Match: 15.55% Indels: 100
DB: 12 Gaps: 21

US-10-047-855-3 (1-672) x US-11-097-143-33685 (1-4100)

QY 14 LeuLeuProGlyGluValPheAlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsn 33
DB 1256 CTGGCCCGGATGTTGGCCCGTGGAGTGGAGATCATCCGGCTCTGGGCTGTGGCAG 1315
QY 34 ProGlnAsnAlaValAlaLeuLeuProGlnAsnAspThrGlyGluSerMetLeuTrpLys 53
DB 1316 CTCGAGAGGCTCTACCCCTCAAGGAGCAGGACAGGCGCCACCAAGTGGTATCTGAGG 1375
QY 54 AlaThrIleValLeuSerArgGlyValSerValGlnTrpArgTyrPheIysGlyTyrPhe 73
DB 1376 GTTGGATATGCGCCAGCCAG-----AGATTCTACTACCGATACCTG 1417
QY 74 LeuGluProLysThrIleGlyGlyProCysGlnValIleValHisLysTrpIleThrHis 93
DB 1418 ATATACTCCAAAGAACAGCAAGGC-----AATCGGGTCTCAGGACCTGGGAGGCCAG 1471
QY 94 LeuGlnProArgSerIleThrProLeuGlu-----SerGluIlele 107
DB 1472 AGAGTGGCTCGATGCTGCAGACCTACGAGATCTACCGCTCCCGGGTCTGGACATCTTC 1531
QY 108 IleAspAspGlyGlnPheGlyIleHisAsnGlyValGluThrLeuAspSerGlyTrpLeu 127
DB 1532 GCGGAGGCTTACCCGCTCTCAGTGGGGGGGATTC---TATCTGGAACGCGTGGTGTG 1588
QY 128 ThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerGluLysProValSerIle 147
DB 1589 CAGTACGAGTACGTGATAGCTAAATTCGCTGGCAGATCATATCTATTTCT--- 1645
QY 148 ThrLysLysLysLeuLysLysSerArgPheArgValLysLeuThrLeuGluGlyLeuGlu 167
DB 1646 -----GGAATGGGAAGTAACGCCAAGTACCGGTTGATCTTACACGCTCAAT 1693
QY 168 GluAspAspAspArgValSerProThrValLeuHisLysMetSerAsnSerLeuGlu 187
DB 1694 TTGATACGACACACGATTT-----GAG 1717
QY 188 IleSerLeuIleSerAspAsnGluPheLysCysArg-His---SerGlnProGluCysG 206
DB 1718 GTCTCCAGATATGCTCAATCATGATCATCTTCGGATTTACCAACCGCTGGACACC-TACAATGC 1777
QY 206 YTyrglyLeuGlnPro-----AspArgTrpThrGluTyr----- 217
DB 1778 TACACACAGGACCATTTGTTGCTTCCTCGGATTTACCAACCGCTGGACACC-TACAATGC 1836
QY 218 -----SerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPhe 234
DB 1837 CCTTCGCTGTGCTATTTATCAGGCATCAAGAGAT---CTCCAGTTA----- 1879
QY 234 eGluGluAspLeuSerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyTh 254
DB 1880 -----TCCCTCGAGAGGCGGTACATATTTCCGGATCAATCAAGGCG----- 1921
QY 254 rAlaCysLeuLeuSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLe 274
DB 1922 -----AGTCGTGGGATTTTGCAGCT 1941
QY 274 uProIleMetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTrpIleI 294

DB 1942 CCCATCTTTTGTAGTTTACAAAAATTTGGCCATTGGGGAGATCACGTTCGCCATATTGGT 2001
QY 294 eIleLysProLeuProGlyTyrSerCys---AspMetLysSerSerPheSerLysTrp 313
DB 2002 GGTTCACCTATGCGCAAAAGTGAGCGCGGAAATCTACAGGCGAGTTTCCATCACTACTG 2061
QY 313 pLysProArgIlePro---LeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrTh 332
DB 2062 GCCGGATAATTGGCCACACATGGATGGATACCGAGGTTTGGAGCCAGCTACTTCCA 2121
QY 332 rAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaSerHi 352
DB 2122 ATCTTCA---ACGAGTCTTACGAGAACACCATCGAAAGTTATCTGGCAGTCTCTGAAGGC 2178
QY 352 sGlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheValProValValTy 372
DB 2179 CAAGGGAGATATGTCACACTGGATGTGCAGCTGACCAAGGACTAGTTCCTCCGTACTG 2238
QY 372 rHisAspLeuThrCysCysLeuThrMetLysLysPheAspAlaAspProValGluLe 392
DB 2239 GCATGGATTGGATTTTACACCTCAGACACAGATCGCTCTGTAAGGATCGCTTGTATT 2298
QY 392 uPheGluIleProValLysGluLeuThrPheAspGlnLeuGlnLeuLeuLysLeuThrHi 412
DB 2299 ACGGTTTGTGTAATAGGGAACCTTACTACTCGGAA-----CTCAAGGCCAGCCG 2349
QY 412 sValThrAlaLeuLysSerLysAspArgLysGluSerValValGlnGlnGluAsnSerPh 432
DB 2350 GGTGTTTATCTCTC-----AAACGTTGACCGTCCAGGAATACACCAATCT 2394
QY 432 eSer-----GluAsnGlnProPheProSerLeuLysMetValLeuG 446
DB 2395 TATGTCAGGATGTGAGCCAAACACAGAAATTTTCCAAAACCTCTCGGAGGTTTGTGA 2454
QY 446 uSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGlnGlnArgAs 466
DB 2455 GGCCCTACCCAAACACTGGGACTTCTGTGGAGATAAAGTGG-----CGCAAAAT-AA 2507
QY 466 pGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMet-AsnLeuPheLeuAspIleI 486
DB 2508 TGGCTTCTGTGTGCCGGAATCCACCCAGAGATTGGAACAAACATTTTACGTGCAGCAA 2567
QY 486 leLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSerPheAspA 506
DB 2568 TTCTACAGATCAGATTCATCATGGGTGGAGCTCCACTAATCTCGCCAGCTTTGATG 2627
QY 506 laAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeuThrG 526
DB 2628 CGGATATATGCAATGATTAGGCTGAAGCAGCATGTCTTCCCTGTGATTCTGATGACA 2687
QY 526 lnglyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThrProIleA 546
DB 2688 TTGGCAAGTCTCAGATTTGGGATGATATGGATCTGAGGCGCACAGAGTTCCAGCAAG 2747
QY 546 laMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGluAspLeuL 566
DB 2748 CATTCAACTTCTTCAGTCCGAGAAATTTGGGAACAGCTCTGCACGTGGAAACTTCC 2807
QY 566 euArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpG 586
DB 2808 AAAACAAACACCCAGCAAGTCAACTTGGCTTGGATCTGCAGCAGTCTCTTTTATATGGG 2867
QY 586 lyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyValAsnGlyL 606
DB 2868 GTAACGATATGCAAGATGAACATCTTCTGGAGCAATTTAGAGCTCTGGATGTACGGGTC 2927
QY 606 euIleTyrAspArgIle-----TyrAspTrpMetProGluGlnProAsnI 621
DB 2928 TGATTTTACATCAGATGATCGAGTGGGTCCGCTTCGCTTGG-----AAACGATCCGAAT 2981
QY 621 lePheGlnValGluGlnLeuGluArgLeu 630
DB 2982 TTTTCCGGCCCCACCAACTTATGGAACATA 3010

Search completed: July 3, 2005, 13:08:50
Job time : 1006 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 3, 2005, 09:24:15 ; Search time 1063 Seconds
(without alignments)
3965.106 Million cell updates/sec

Title: US-10-047-855-3
Perfect score: 3537
Sequence: 1 MTPSQVAFIRGTLLPGEVF.....LCGESDIHVANGINDVENA 672

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q=/cgn2_1/USPTO.spool/US10047855/runat_01072005_155531_2380/app_query.fasta_1.839
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR WAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10047855@cgn 1 1 740 runat_01072005_155531_2380
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US16_NEW_PUB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3537	100.0	2019	21	US-10-772-636-25	Sequence 25, Appl
2	3537	100.0	3206	16	US-10-047-855-4	Sequence 4, Appl
3	3537	100.0	3206	17	US-10-426-776-34	Sequence 34, Appl
4	3537	100.0	3265	21	US-10-887-553A-1007	Sequence 1007, Ap
5	3537	100.0	3499	10	US-09-814-353-21000	Sequence 21000, A
6	3537	100.0	3499	20	US-10-357-930-23730	Sequence 23730, A
7	3537	100.0	3499	20	US-10-357-930-23632	Sequence 29632, A
8	2965.5	83.8	3381	17	US-10-426-776-39	Sequence 39, Appl
9	2552	72.2	2738	17	US-10-426-776-32	Sequence 32, Appl
10	1399	39.6	2393	17	US-10-426-776-48	Sequence 48, Appl
11	814	23.3	471	9	US-09-998-598-1849	Sequence 1849, Ap
12	717	20.3	520	10	US-09-814-353-16669	Sequence 16669, A
13	698	19.7	436	9	US-09-796-692-8067	Sequence 8067, Ap
14	698	19.7	436	14	US-10-040-862-8067	Sequence 8067, Ap
15	698	19.7	436	17	US-10-057-475B-8067	Sequence 8067, Ap
16	698	19.7	436	17	US-10-154-884B-8067	Sequence 8067, Ap
17	698	19.7	436	19	US-10-764-324-8067	Sequence 8067, Ap
18	628	17.8	396	10	US-09-814-353-3977	Sequence 3977, Ap
19	628	17.8	396	10	US-09-814-353-10285	Sequence 10285, A
20	570	16.1	425	20	US-10-357-930-35883	Sequence 35883, A
21	570	16.1	425	20	US-10-357-930-4846	Sequence 4846, A
22	564	15.9	327	20	US-10-357-930-15054	Sequence 15054, A
23	482.5	13.6	3672	17	US-10-369-493-46175	Sequence 46175, A
24	472.5	13.4	359	20	US-10-357-930-5885	Sequence 5885, A
25	432	12.2	3799	17	US-10-369-493-26821	Sequence 26821, A
26	430.5	12.2	2226	17	US-10-369-493-36472	Sequence 36472, A
27	347	9.8	193	9	US-09-864-761-31597	Sequence 31997, A
28	343	9.7	463	9	US-09-864-761-2380	Sequence 2380, Ap
29	339	9.3	1121	14	US-10-198-846-11446	Sequence 11446, A
30	304.5	8.6	1540	20	US-10-425-115-168092	Sequence 168092, A
31	299	8.5	1499	19	US-10-437-963-16400	Sequence 16400, A
32	298.5	8.4	1410	18	US-10-425-114-27639	Sequence 27639, A
33	298	8.4	1391	18	US-10-425-114-22987	Sequence 22987, A
34	291	8.2	401	9	US-09-864-761-15483	Sequence 15483, A
35	286.5	8.1	1596	18	US-10-425-114-31409	Sequence 31409, A
36	286.5	8.1	1724	20	US-10-425-115-162869	Sequence 162869, A
37	282	8.0	1399	18	US-10-425-114-22946	Sequence 22946, A
38	282	8.0	1637	20	US-10-425-115-168094	Sequence 168094, A
39	273	7.7	1523	18	US-10-425-114-22914	Sequence 22914, A
40	273	7.7	1604	20	US-10-425-115-34867	Sequence 34867, A
41	272.5	7.7	1949	20	US-10-739-930-755	Sequence 755, App
42	271.5	7.7	1164	19	US-10-437-963-96985	Sequence 96985, A
43	259	7.3	323	9	US-09-728-445-622	Sequence 622, App
44	255	7.2	151	9	US-09-864-761-19114	Sequence 19114, A
45	251.5	7.1	1191	19	US-10-437-963-22646	Sequence 22646, A

ALIGNMENTS

RESULT 1
US-10-772-636-25
; Sequence 25, Application US/10772636
; Publication No. US20050042687A1
; GENERAL INFORMATION:
; APPLICANT: Kelly, Louise M.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Farlow, Deborah
; APPLICANT: Healy, Aileen
; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,
; TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,
; TITLE OF INVENTION: 9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,
; TITLE OF INVENTION: 9137, 13908, 14310, 17600, 23584, 27824, 28469, 38947,
; TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR
; TITLE OF INVENTION: 5014

```

; FILE REFERENCE: MPI03-015PIRNOWNMIM
; CURRENT APPLICATION NUMBER: US/10/772,636
; CURRENT FILING DATE: 2004-02-05
; PRIOR APPLICATION NUMBER: US 60/445,241
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/448,389
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/456,320
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/460,279
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/465,924
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: US 60/470,052
; PRIOR FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: US 60/498,106
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 60/500,179
; PRIOR FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: US 60/502,909
; PRIOR FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: US 60/510,351
; PRIOR FILING DATE: 2003-10-10
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2019)
US-10-772-636-25

Alignment Scores:
Pred. No.: 0 Length: 2019
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-10-047-855-3 (1-672) x US-10-772-636-25 (1-2019)

QY 1 MetThrProSerGlnValAlaPheGluLeuArgGlyThrLeuLeuProGlyGluValPhe 20
DB 1 ATGACACCTTCTCAGGTGGCTTTGAAATAGAGGAACTCTTTTACCAGGAGAAATTTT 60

QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAanTrpAsnProGlnAsnAlaValAlaLeu 40
DB 61 GCGATATGTGGAAGCTGTGATGCTTTGGAAACTCGAAATCTCTCAAAATGCTGTGCTCTT 120

QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTlpLysAlaThrIleValLeuSerArg 60
DB 121 CTTCCAGAGAAATGACACAGGTGAAGCATGCTATGGAAAGCAACCATTTGACTCAGTAGA 180

QY 61 GlyValSerValcInltnrYzArgTyrPhetySgGlyTyrPheLeuGluProLysThrIleGly 80
DB 181 GGAGTATCAGTTTCAGTATCGTACTTCAAAGGGTACTTTTATAGAACCAAGACTATCGT 240

QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
DB 241 GGTCCATGTCAAGTGTAGTAGTTCCAAAGTGGGAGACTCATCTCAACACCGATCAATAACC 300

QY 101 ProLeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
DB 301 CCTTTTGAAGCGAAATATTATTGACGATGGCAATTTGGAAATCCCAATGGTGTGTGAA 360

QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgIleuHisTyrSer 140
DB 361 ACTCTGGATTCGATGGCTGCACATGTGCAGACTGAAATAAGATTACGTTTGCAATTATCT 420

```

```
Db 1501 TTTTCTTCATTTGATCGATATTTTCACAAATGGTTTCGCAAAAGCAGAACAAATATCCG 1560
QY 521 IleLeuPheLeuThrGlnGlySerGluIleTyrProGluLeuMetAspLeuArgSer 540
Db 1561 ATACTATTTTAACTCAAGGAAAATCTGAGATTATCTCTGAACTCATGGACCTCAGATCT 1620
QY 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
Db 1621 CGACAAACCCCATTCGAATGAGCTTTGCACAGTTTGAAGAAATCTACTGGGGATAAATGTA 1680
QY 561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaIlysAlaIysGlyLeu 580
Db 1681 CATACTGAAGACTTGCTCAGAAACCCATCTATATTCAAGAGCAAAAGCTAAGGGACTA 1740
QY 581 ValIlePheCysTrpGlyAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
Db 1741 GTCATATTCTGTGGGGTGATGATACCAATGATCTCTGAAAACAGAGGAATTTGAAGGA 1800
QY 601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
Db 1801 CTTGGAGTTAATGGTCTAATTATGATAGATATATGATGGATGCTGAAACCAAT 1860
QY 621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
Db 1861 ATATTCCAAAGTGGAGCAATTGGAACGCTTGAAGCAGGAATTTGCCAGAGCTTAAGAGCTGT 1920
QY 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 1921 TTGTGTCCCACTGTAGCCGCTTTGTTCCTCATCTTTGTGTGGGGAGTCTGATATCCAT 1980
QY 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 1981 GTGGATGCCAACGGCATTGATAACGTGGAGAAATGCT 2016
```

RESULT 2

```
US-10-047-855-4
; Sequence 4, Application US/10047855
; Publication No. US20030165863A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Lillian Wei-Ming
; TITLE OF INVENTION: NARC10 and NARC16, Programmed Cell
; FILE REFERENCE: Death-Associated Molecules and Uses Thereof
; FILE REFERENCE: 35800/242056
; CURRENT APPLICATION NUMBER: US/10/047,855
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/262,306
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3206
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145)...(2163)
; NAME/KEY: misc.feature
; LOCATION: (1)...(17)
; OTHER INFORMATION: Vector sequence
US-10-047-855-4
```

```
Alignment Scores:
Pred. No.: 0 Length: 3206
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0
```

US-10-047-855-3 (1-672) x US-10-047-855-4 (1-3206)

QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20

```
Db 145 ATGACACCTTTCAGGTTGCCTTTTGAAATAAGAGGAACCTCTTTTACCAGGAGAGTTTTT 204
QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 205 GCGATATGTGGAGCTGTGATGCTTTTGGGAAACTGGAACTCTCAAAATGCTGTGGCTCTT 264
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 265 CTTCCAGAGAAATGACACAGGTGAAAGCATGCTATGGAAGCAACCACTTGTACTCAGTAGA 324
QY 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 325 GGAATATCAGTTTCAGTATCGCTACTTCAAAAGGGTACTTTTATAGAACCAAAAGACTATCGGT 384
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 385 GGTCCATGTCAAGTGTAGTTCACAAGTGGGAGACTCATCTACACCACCATCAATAACC 444
QY 101 ProLeuGluSerGluIleIleLeuAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
Db 445 CCTTTAGAAAAGCGAAATTAATTATTCAGCATGACAAATTTGGAATCCACAAATGGTGTGAA 504
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
Db 505 ACTCTGGATTTCTGGATGGCTGACATGTCTCAGACTGAAATTAAGATTAAGTTTGCATATTCT 564
QY 141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
Db 565 GAAAACCTCTCTGTGTCAATAACCAAGAAAATAATTAATAAATAATCTAGATTAGGGTGAAG 624
QY 161 LeuThrLeuGluGlyLeuGluAspAspAspArgValSerProThrValLeuHis 180
Db 625 CTGACACTAGAAAGGCTGGAGGAAGATGACGATGATAGGGTATCTCCCACTGTACTCCAC 684
QY 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
Db 685 AAAATGTCCAATAGCTTGGAGATATCTTTAATAAGCGCAATAGAGTTCAAGTGCAAGGCAT 744
QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
Db 745 TCACAGCCGAGTGTGGTTATGGCTTCGAGCTGATCGTTGGACAGAGTACAGCATACAG 804
QY 221 ThrMetGluProAspAsnLeuGluIlePheAspPhePheGluGluAspLeuSerGlu 240
Db 805 ACGATGGAACCAAGATAAACCCTGCAACTAATCTTTTGAATTTTTCGAAGAAGATCTCAGTGAG 864
QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db 865 CACGTAGTTTCAGGGTGATGCCCTTCTCTGGACATGTGGGTACAGCTTGTCTTATCATCC 924
QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
Db 925 ACCATTGTCTGAGAGTGGAAAGAGTGTGGAATCTTACTCTTCCCATCATGAGCAAAAT 984
QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGly 300
Db 985 TCCCGGAAAAACAATAGGCAAGTGAGAGTTGACTATATAATTTATTAAGCCATTACCAGGA 1044
QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
Db 1045 TACAGTTGTGACATGAAATCTTTCATTTTCCAAAGTATTGGAAGCCCAAGAAATACCATTTGGAT 1104
QY 321 ValGlyHisArgGlyValAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
Db 1105 GTTGCCCATCGAGGTGCGGAAACTCTACAAACATGCCAGCTGGCTAAGATTCAAGAA 1164
QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyValAlaPheValGluPheAsp 360
Db 1165 AATACTATTGCTTCTTTAAGAAATGCTGCTAGTCTCATGTGTGCAGCCCTTCTAGAAATTGAC 1224
QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
```



```

1225 GTACACCTTTCAAAGGACTTTGTGCCCGTGTGTATATCATGATCTTACCTGTTGTTGACT 1284
Qy
381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluLeuProValLysGluLeu 400
Db
1285 ATGAAAAGAAATTTGATGCTGATCCAGTTCGAATTAATTTGAAATTTCCAGTAAAGAAATTA 1344
Qy
401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
Db
1345 ACATTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCAGCTGAATCTTAAGAT 1404
Qy
421 ArgLysGluSerValValGlnGluLeuAsnSerPheSerGluAsnGlnProPheProSer 440
Db
1405 CGAAAGAAATCTGTGTTGAGAGGAAATTTCTTTTCAGAAATTCAGCAATTTCTTCT 1464
Qy
441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluLeuLysTrp 460
Db
1465 CTTAAGATGTTTATAGATCTTTGCCAGAGATGTAGGTTTAAACATTTGAATTAATGG 1524
Qy
461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
Db
1525 ATCTGCCAGCAAGGATGGAATGTGGATGGTAACTTATCAACATATTTTGACATGAAT 1584
Qy
481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal 500
Db
1585 CTGTTTTCGATATATTTTAAAACTGTTTGTAGAAATTTCTGGAGAGGAGATAGT 1644
Qy
501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTrpPro 520
Db
1645 TTTTCTTCATTTGATGCAGATTTTGCACATGGTTTCGCAAAAGCAGCAAAATATCCG 1704
Qy
521 IleLeuPheLeuThrGlnGlyLysSerGluLeuIleTyrProGluLeuMetAspLeuArgSer 540
Db
1705 ATACTATTTTAACTCAAGCAAAATCTGAGATTTATCTTGAATCTCATGGACCTCAGATCT 1764
Qy
541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
Db
1765 CGGACAAACCCCAATTCGAATGAGCTTTGCACAGTTTGAATAATCTACTGGGGATTAATGTA 1824
Qy
561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysLysGlyLeu 580
Db
1825 CATACTGAAGACTTGTCTAGAAACCCATCTTATATTCAGAGGCAAAAGCTAAGGACTA 1884
Qy
581 ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLysGlu 600
Db
1885 GTCATATTTCTGCTGGGTGATGATACCAATGATCTCTGAAACAGAGAAATTTGAAGGAA 1944
Qy
601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
Db
1945 CTTGGAGTTAATGGTCTAAATTTATGATAGGATATATGATTTGGATGCTGACCAACCAAT 2004
Qy
621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
Db
2005 ATATTCCAAGTGGAGCAATTTGGAACGCTGAAGCAGGAATTCGACAGCTTAAAGACTGT 2064
Qy
641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db
2065 TTGTGTCCTACTGTTAGCCGCTTGTTCCTCATCTTTGTGTGGGGAGTCTGATATCCAT 2124
Qy
661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db
2125 GTGGATGCCAACGGCATTTGATAACGTGGAGAAATGCT 2160

```

RESULT 3

US-10-426-776-34

```

; Sequence 34, Application US/10426776
; Publication No. US20040009553A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Williamson, Mark J.
; APPLICANT: Tsia, Fong-Ying
; APPLICANT: Rudolf-Owen, Laura A.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Meyers, Rachel E.

```

```

; APPLICANT: Chiang, Lillian Wei-Ming
; APPLICANT: Hunter, John Joseph
; APPLICANT: Wood, Andrew
; APPLICANT: Jenkins, Lorayne P.
; TITLE OF INVENTION: NOVEL 27411, 23413, 22438, 23553.
; TITLE OF INVENTION: 25278, 26212, NARC SCI, NARC 10A, NARC 1, NARC 12, NARC 13,
; TITLE OF INVENTION: NARC17, NARC 25, NARC 3, NARC 4, NARC 7, NARC 8, NARC 11,
; TITLE OF INVENTION: NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20, NARC 26, NARC
; TITLE OF INVENTION: 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C, NARC
; TITLE OF INVENTION: 8B, NARC 9, NARC2A, NARC 16B, NARC 1C, NARC 1A, NARC 25,
; TITLE OF INVENTION: 86604 AND 32222 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MP103-0620NMIM
; CURRENT APPLICATION NUMBER: US/10/426, 776
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 10/229, 662
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/795, 691
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185, 517
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 10/105, 992
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/406, 045
; PRIOR FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 10/314, 881
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 09/773, 426
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 09/495, 823
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/692, 785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/161, 188
; PRIOR FILING DATE: 1999-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 3206
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-10-426-776-34

```

```

Alignment Scores:
Pred. No.: 0 Length: 3206
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

```

US-10-047-855-3 (1-672) x US-10-426-776-34 (1-3206)

```

Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 145 ATGACACCTTCTCAGGTTGCCTTTGAAATAAGAGGAACCTCTTTTACCAGGAGAAAGTTTTT 204
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 205 GCGATATGTGGAAGCTGTGATGCTTTGGAAATACTGGAATCTCAAAATGCTGTGGCTCTT 264
Qy 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 265 CTTCCAGAGATGACACAGGTGAAAGCATGCTATGGAAGCAACCATTTGTACTCATAGTA 324
Qy 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 325 GGAGTATCAGTTTCAGTATCGCTACTTTCAAAGGGTACTTTTAAAGAACCAAGAACTATCGGT 384
Qy 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 385 GTTCCATGTCAAGTGATAGTTTCAAGTGGGAGACTCTCTACACACCAGCATCAATAACC 444

```

101 ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
141 GluLysProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
161 LeuThrLeuGluLysLeuGluAspAspAspArgValSerProThrValLeuHis 180
181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
221 ThrMetGluProAspAsnLeuGluIlePheAspPheGluGluAspLeuSerGlu 240
241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300
301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysLysLeuThr 380
381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
401 ThrPheAspGlnLeuGlnLeuLeuLysLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
421 ArgLysGluSerValValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSer 440
441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480

1525 ATCTGCCACCAAGGATGGAATGGGGATGGTAATCTTATCAACATATTTTGCATGAAT 1584
481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal 500
1585 CTGTTTTCGATATAATTTTAAACTGTTTAAAGAAATCTGGGAAGAGGAGATAGTG 1644
501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro 520
1645 TTTTCTTCATTTGATGCAGATATTTGCACAATGGTTTCGCAAAAGCAGAAATATCCG 1704
521 IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer 540
1705 ATACTATTTTAACTCAAGCAAACTCTGAGATTTATCTCGAACTCATGGACCTCAGATCT 1764
541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
1765 CGGACACACCCCATTTGCAATGAGCTTTGCACAGTTTGAANAATCTACTGGGGATAAATGTA 1824
561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
1825 CATACTGAAGACTTGTCTCAGAAACCATCTATATTCAAGAGGCAAAAGCTAAGGACTA 1884
581 ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
1885 GTCATATTCTGTGGGGTGATGATACCAATGATCTCTGAAAGCAGAAAGAAATTTGAAGGAA 1944
601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
1945 CTGGAGTTAATGGTCTAATTTATGATAGATATATGATGATGCTGGAACACCAAT 2004
621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnLeuProGluLeuLysSerCys 640
2005 ATATTCCAAAGTCGAGCAATTGGAACGCTGGAAGCAGCAATTTGCCAGAGCTTAAGAGCTGT 2064
641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
2065 TTGTGTCCCACTGTTAGCCGCTTTGTTCCTCATCTTTGTGTGGGGAGTCTGATATCCAT 2124
661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
2125 GTGGATGCCACGGCATTTGTAACGTGGAGATGCT 2160

RESULT 4

US-10-887-553A-1007
; Sequence 1007, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; TITLE OF INVENTION: with insulin signalling dysregulation
; FILE REFERENCE: 4-33262
; CURRENT APPLICATION NUMBER: US/10/887,553A
; PRIOR FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; PRIOR FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1007
; LENGTH: 3265
; TYPE: DNA
; ORGANISM: human
US-10-887-553A-1007

Alignment Scores:

Pred. No.: 0 Length: 3265
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 21
DB: 0

US-10-047-855-3 (1-672) x US-10-887-553A-1007 (1-3265)

Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 213 ATGACACCTTCTCAGGTGGCTTTGAAATAAGAGAACTCTTTTACCAGGAGAGTTTTT 272

Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 273 GCGATATGTGGAGCTGTGATGCTTTGGGAACTGGAACTCTCAAAATGCTGTGGCTCTT 332

Qy 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 333 CTTCCAGAGAATGACACAGGTGAAGCATGTATGGAAGCAACCAATGTACTCAGTAGA 392

Qy 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 393 GCGATGATCAGTTCACTATCGCTACTTCAAGGGTACTTTTTAGAACCCAAAGACTATCGT 452

Qy 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 453 GGTCCATGTCAAGTGTAGTTCACAAAGTGGAGACTCATCTACAAACCAACGATCAATAACC 512

Qy 101 ProLeuGluSerGluIleIleLeuAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
Db 513 CCTTTAGAAAGCGAAATATTATTGACGATGGACAATTTGGAAATCCAAATGGTGTGAA 572

Qy 121 ThrLeuAspSerGlyTripleThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
Db 573 ACTCTGGATTCTGGATGGCTGATGTCTCAGACTGAAATTAAGAAATACGTTTGCATTTCT 632

Qy 141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
Db 633 GAAAAACCTCTCTGTGTCAATATACCAAGAAATAAATAAATAATCTAGATTTAGGGTGAAG 692

Qy 161 LeuThrLeuGluGlyLeuGluAspAspArgValSerProThrValLeuHis 180
Db 693 CTGACACTAGAACCCCTGGAGGAAGATGACGATGATAGGGTATCTCCCACTGTACTCCAC 752

Qy 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
Db 753 AAAATGTCCAATAGCTTGGAGATATCTTAATAAGCGACAATGAGTTCAAGTGCAGGAT 812

Qy 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
Db 813 TCACAGCCGAGTGTGGTATGCTTGCAGCCGTGATCGTTGCAGAGAGTACAGCATACAG 872

Qy 221 ThrMetGluProAsnLeuGluLeuIlePheAspPhePheGluAspLeuSerGlu 240
Db 873 ACCATGGAACCAATAGCTTGAACCTTGAATCTTTTGAATTTTTCGAAGAAGATCTCAGTGAG 932

Qy 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db 933 CACGTAGTTCCAGGGTGTATGCTTCTCGGACATGTGGGTACAGCTGTCTTTATCATCC 992

Qy 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
Db 993 ACCATTCTCAGAGTGGAAAGTGTGGAAATCTTACTCTTCCCATCATGACAGCAAT 1052

Qy 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300
Db 1053 TCCCGGAACAATAGCCAAAGTGAGAGTTGACTATATAATTATTAAAGCCATTACCAGGA 1112

Qy 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
Db 1113 TACAGTTGTGACATGAAATCTTCAATTTTCCAAGTATTGGAAGCCAAAGATACCATTGGAT 1172

Qy 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
Db 1173 GTTGGCATTCAGGTGACGAGAACTCTACAAACACTGCCAGCTGGCTTAAGTCAAGAA 1232

Qy 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
Db 1233 AATACTATTGCTTCTTTAAGAAATGCTGCTAGTCAATGTCAGCCCTTTGTAGAAATTTGAC 1292

Qy 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
Db 1293 GTACACCTTTCAAAGGACTTTGTGCGGTGGTATATCATGATCTTACTGTGTGTTGACT 1352

Qy 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
Db 1353 ATGAAATAAGAAATTTGATGCTGATCCAGTTGAATTTATTGAAATTTCCAGTAAAGAAATTA 1412

Qy 401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
Db 1413 ACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCATGAAATCTAAGGAT 1472

Qy 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
Db 1473 CGGAAAGAAATCTGTGTTTCCAGGAGGAAATTTCTTTTTCAGAAATTCAGGCATTTCTCTTCT 1532

Qy 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
Db 1533 CTTAAGATGGTTTAGAGTCTTTTGCAGAAAGATGTAGGGTTTAAACATTGAAATAAATG 1592

Qy 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
Db 1593 ATCTGCCAGCAAGGATGGAATGTGGATGGTAACTTATCAACATATTTTGACATGAAT 1652

Qy 481 LeuPheLeuAspIleIleLysThrValLeuGluAsnSerGlyLysArgArgIleVal 500
Db 1653 CTGTTTTTGGATATAATTTTAAAACTGTTTAAAGAAATTTCTGGGAGAGGAGAAATAGTG 1712

Qy 501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTrpPro 520
Db 1713 TTTCTTTCAITTTGATGACGATATTTGCACAACTGGTTCCGCAAGCAGAACTAATATCCG 1772

Qy 521 IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer 540
Db 1773 ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTCAACTCATGGACCTCAGATCT 1832

Qy 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal 560
Db 1833 CGGCAACCCCCCAATGCAATGAGCTTTGCACAGTTTGGAAATCTACTGGGGATATAATGTA 1892

Qy 561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
Db 1893 CATACTGAAGACTTGTCTCAGAAACCATCTATATTTCAAGAGGCAAAAGCTAAGGGACTA 1952

Qy 581 ValllePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
Db 1953 GTCATATTTCTGCTGGGTGATGATACCAATGATCTCTGAAACAGAGGAAATTTGAAGGAA 2012

Qy 601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
Db 2013 CTTGGAGTTAATGGTCTAAATTTATGATAGGATATATGATGGATGCTTGAACACCAAT 2072

Qy 621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
Db 2073 ATATTCCAAAGTGGAGCAATTTGGAACGCTGAAGCAGGAAATTTGCCAGAGCTTAAGAGCTGT 2132

Qy 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 2133 TTGTGTCACACTGTTTAGCCGCTTTGTCCTCATCTTTGTGTGGGGAGTCTGATATCCAT 2192

Qy 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2193 GTGGATGCCAAGCGCATTTGATAACGTGGAGAAATGCT 2228

RESULT 5
US-09-814-353-21000
; Sequence 21000, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James

```

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIORITY APPLICATION NUMBER: US 60/191,031
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY APPLICATION NUMBER: US 60/207,124
; PRIORITY FILING DATE: 2000-05-25
; PRIORITY APPLICATION NUMBER: US 60/211,940
; PRIORITY FILING DATE: 2000-06-15
; PRIORITY APPLICATION NUMBER: US 60/216,820
; PRIORITY FILING DATE: 2000-07-07
; PRIORITY APPLICATION NUMBER: US 60/220,661
; PRIORITY FILING DATE: 2000-07-25
; PRIORITY APPLICATION NUMBER: US 60/257,672
; PRIORITY FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21000
; LENGTH: 3499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3490, 3491, 3492, 3493, 3494, 3495, 3496, 3497, 3498, 3499
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21000

```

Alignment Scores:

```

Pred. No.: 0 Length: 3499
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

```

US-10-047-855-3 (1-672) x US-09-814-353-21000 (1-3499)

```

QY 1 MetThrProSerGlnValAlaPheGluLeuArgGlyThrLeuLeuProGlyGluValPhe 20
DB 221 ATGACACCTTCTCAGGTTGCTTGTGAATGAAGAGAACTCTTTTACGAGAGAGTGT 280
QY 21 AlaileCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
DB 281 GCGATATGTGAAGCTGTGATGCTTTGGGAACTGGATCTCTCAAAATGCTGGCTCT 340
QY 41 LeuProGlnAsnAspThrGlyGluSerMetLeuTrpLysAlaThrileValLeuSerArg 60
DB 341 CTTCCAGAGAATGACACAGGTGAAGCATGCTATGGAAGCAACCATTTGACTCAGTAGA 400
QY 61 GlyValSerValGlnTrpArgGlyPheLysGlyTyrPheLeuGluProLysThrileGly 80
DB 401 GGAGTATCAGTTCAGTATCCCTACTCTCAAGGGGTACTTTTTAGAACCAAGACTATCGGT 460
QY 81 GlyProCysGlnValileValHisLysTrpGluThrHisLeuGlnProArgSerilleThr 100
DB 461 GGTCCATGTCAGTGATAGTTTCAACAGTGGAGACTCATCTCAACCCAGTCAATTAACC 520
QY 101 ProLeuGlnSerGluilellelelelelelelelelelelelelelelelelelelele 120
DB 521 CCTTTAGAAAGCGAAATTTATTTATGACATGGACAAATTTGGAATCCACATGCTGTGAA 580
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluilelelelelelelelelelelele 140
DB 581 ACTCTGATTTCTGGATGGCTGACATGTACAGCTGAATTAAGATTTAGCTTTGCAATATCT 640
QY 141 GluLysProProValSerilleThrLysLysLysLysLysLysSerArgPheArgValLys 160
DB 641 GAAATACCTCTCTGTCTCATTAACCAAGAAAATTAATAAATCTAGATTTAGGTGGAAG 700
QY 161 LeuThrLeuGluGlyLeuGluAspAspArgValSerProThrValLeuHis 180

```

```

DB 701 CTGACACTAGAGGCTTGAGGAAGATGACGATAGGATATCTCCCACTGACTCCAC 760
QY 181 LysMetSerAsnSerLeuGluilelelelelelelelelelelelelelelelelelelele 200
DB 761 AAAATGTCCATAGCTTGGAGATATCTTAAATGAAGCAATGAGTTCAAGTCAGGCAT 820
QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTrpSerilleGln 220
DB 821 TCACAGCGGAGTGTGTTATGGCTTTCAGCCCTGATCGTTGGACAGAGTACAGATACAG 880
QY 221 ThrMetGluProAspAsnLeuGluilelelelelelelelelelelelelelelelelele 240
DB 881 ACGATGGAACCAAGATAACCTGGAACATAATCTTTGATTTTTCGAAGAAGATCTCAGTGAG 940
QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
DB 941 CAGGTAGTTTCAGGCTGATGCCCTTCTCGACATGTGGGTACAGCTTGTCTCTTATCATCC 1000
QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
DB 1001 ACCATTGCTGAGAGTGGAAAGAGTGTGGAATTTCTTACTCTTCCCATCATGAGCAAAAT 1060
QY 281 SerArgLysThrileGlyLysValArgValAspTyrIlelelelelelelelelelelele 300
DB 1061 TCCCGGAAACCAATAGCAAGAGTGAAGTTGACTATATAATTAATTAAGCCATTTACCAGGA 1120
QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgilleProLeuAsp 320
DB 1121 TACAGTTGTGACATGAATCTTTCATTTCCAGTATTTGGAGCCAGAGATACCATTTGGAT 1180
QY 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
DB 1181 GTTGGCCATCGAGGTGCAGGAAACTCTACAACAATGCCAGCTGGCTAAAGTTCAAGAA 1240
QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
DB 1241 AATACTATTGCTTCTTTAAGAAATGCTGCTAGTCATGGTGCAGCCCTTTGAGAAATTTGAC 1300
QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
DB 1301 GTACACCTTTCAAGAGACTTTGTGCCGTGGTATATCATGATCTTACCTGTGTTGACT 1360
QY 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluileProValLysGluLeu 400
DB 1361 ATGAAAGAAATTTGATGCTGATCCAGTTGAATTTGAAATTTGAAATTCAGTAAAGAAATTA 1420
QY 401 ThrPheAspGlnLeuGlnlelelelelelelelelelelelelelelelelelelelele 420
DB 1421 ACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTGAAATCTAAGGAT 1480
QY 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
DB 1481 CGGAAGAATCTGTGGTTGAGGAGGAAATTCCTTTTCAGAAAATCAGCCATTTCTCTCT 1540
QY 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnlelelelelelelele 460
DB 1541 CTTAAGATGGTTTAGAGTCTTTGCCAGAAGATGTAGGGTTTAAACATTGAATAATAATGG 1600
QY 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrThrPheAspMetAsn 480
DB 1601 ATCTGCGACCAAGGAGGATGGAATGTGGGATGGTAACTTATCAACATATTTTGCATGAAT 1660
QY 481 LeuPheLeuAspIlelelelelelelelelelelelelelelelelelelelelelele 500
DB 1661 CTGTTTTTGGATATAATTTTAAAAACTGTTTTAGAAAAATCTCGGAGAGGAGAAATAGTG 1720
QY 501 PheSerSerPheAspAlaAspilleCysThrMetValArgGlnLysGlnAsnLysTrpPro 520
DB 1721 TTTTCTTCTTATGATGCAGATATTTGCACAAATGGTTTCGCAAAAGCAGAAACAAATATCCG 1780
QY 521 IleLeuPheLeuThrGlnGlyLysSerGluilelelelelelelelelelelelelelelele 540

```

1781 ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTCTGAACCTCATGGACCTCAGATCT 1840
Db
541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAenLeuLeuGlyLeuVal 560
Qy
1841 CGGACACCCCATTTGCAATGAGCTTTGCACAGTTTGAATAATCTACTGGGATAAATGTA 1900
Db
561 HisThrGluAenLeuLeuArgAenProSerTyrIleGlnGluAlaLeuAlaLeuGlyLeu 580
Qy
1901 CATATCTGAAGACTTCTCAGAACCCATCCCTATATTAAGAGGGCAAGAGCTTAAGGACTA 1960
Db
581 ValIlePheCysTyrGlyVasPAspThrAsnAspProGluAenArgArgLeuLeuGlyGlu 600
Qy
1961 GTCATATCTCTGGGTGATCATCAATGATCTGTAAGAGGCAAGAAATTTGAAGGAA 2020
Db
601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTyrMetProGluGlnProAsn 620
Qy
2021 CTTGAGTTAATGGTCTAATTTATGATAGGATATGATGATGATGCTTGAACCAACCAAT 2080
Db
621 IlePheGlnValGluGlnLeuGluArgLeuGlyGlnGluLeuProGluLeuLeuSerCys 640
Qy
2081 ATATTCCAAAGTGAGCAATTTGAACGCCCTGAAAGCAGGAATGCCAGAGCTTAAGAGCTGT 2140
Db
641 LeuCysProThrValSerArgPheValProSerSerLeuLeuGlyGluSerAspIleHis 660
Qy
2141 TTGTGTCCTCACTGTAGCGCTTTGTTCCCTCATCTTTGTGTGGGAGTCTGATATCCAT 2200
Db
661 ValAspAlaAsnGlyIleAspAsnValGluAenAla 672
Qy
2201 GTGGATGCCAACGGCATTGATAACGTGGAGATGCT 2236
Db

RESULT 6

US-10-357-930-23730
; Sequence 23730, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23730
; LENGTH: 3499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 3490, 3491, 3492, 3493, 3494, 3495, 3496, 3497, 3498, 3499
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23730

Alignment Scores:
Pred. No.: 0 Length: 3499
Score: 3537.00 Matches: 672

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0
US-10-047-855-3 (1-672) x US-10-357-930-23730 (1-3499)
Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 221 ATGACACCTTCTCAGGTGGCTTTGAAATTAAGAGAACTCTTTTACCAGAGAGATTTTT 280
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTyrAsnProGlnAenAlaValAlaLeu 40
Db 281 GCGATATGTGGAAGCTGTGATGCTTTGGGAACTGGAATCTCTCAAAATGCTGTGGCTCTT 340
Qy 41 LeuProGluAenAspThrGlyGluSerMetLeuTyrPheAlaThrIleValLeuSerArg 60
Db 341 CTTCCAGAGAAATGACACAGGTGAAAGCATGCTATGGAAGCAACCATTTGTACTCAGTAGA 400
Qy 61 GlyValSerValGlnTyrArgTyrPheLeuGlyTyrPheLeuGluProLysThrIleGly 80
Db 401 GGAGTATCAGTTCAGTATCGTACTTCAAAGGGTACTTTTATAGAACCAAGACTATCGGT 460
Qy 81 GlyProCysGlnValIleValHisLysTyrGluThrHisLeuGlnProArgSerIleThr 100
Db 461 GGTCCATGTCAAGTGTAGTTCACAGTGGGAGACTCATCTACAAACCAACCATCAATAACC 520
Qy 101 ProLeuGluSerGluIleIleLeuAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
Db 521 CCITTTAGAAAGCGAAATTTATTTGACGATGGACAAATTTGGAATCCACATGGTGTGAA 580
Qy 121 ThrLeuAspSerGlyTyrLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
Db 581 ACTCTGGATTTCTGGATGGCTGACATGTCAGACTGAAATAGATTACGTTTGCATTATCT 640
Qy 141 GluLysProValSerIleThrLysLysLysLysLysSerArgPheArgValLys 160
Db 641 GAAAAACCTCTGCTGTCAATAACCAAGAAAAAATAAAAAAATCTAGATTTAGGGTGAAG 700
Qy 161 LeuThrLeuGlyLeuGluGluAspAspAspArgValSerProThrValLeuHis 180
Db 701 CTGACACTAGAAAGCGCTTGGAGGAAGATGACATGATAGGGTATCTCCACCTGTACTCCAC 760
Qy 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
Db 761 AAAATGTCATAGCTTGGAGATATCTTTATAGCGCAATAGTTCAAGTCAGGCAT 820
Qy 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTyrThrGluTyrSerIleGln 220
Db 821 TCACAGCCGGAGTGTGGTTATGGCTTGACGCTGATCGTTGGACAGAGTACAGCATACAG 880
Qy 221 ThrMetGluProAspAsnLeuLeuIlePheAspPhePheGluGluAspLeuSerGlu 240
Db 881 ACGATGGAAACCAAGATAACCTGGAACATACTTTTGAATTTTTCGAAGAAGATCTCAGTGAG 940
Qy 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db 941 CACGTAGTTTCAGGGTGAAGCCCTTCTCGACATGTTGGGTACAGCTGTCTCTTATCATCC 1000
Qy 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAen 280
Db 1001 ACCATTGCTGAGAGTGGAAAGAGTCTGGAATTTCTTACTCTTCCCATCATGAGCAAAAT 1060
Qy 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGly 300
Db 1061 TCCCGGAAACCAATAGGCAAGTGAGAGTTGACTATATAATTAATTAAGCCATTACAGGA 1120
Qy 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTyrLysProArgIleProLeuAsp 320
Db 1121 TACAGTTGTGACATGAATCTTCATTTCCAGATATTGGAAGCCAGAAATACCATTTGGAT 1180
Qy 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340

```
Db 1181 GTTGGCCATCGAGTCCAGAAACTCTACAACAACCTGCCAGCTGGCTAAAGTTCAAGAA 1240
Qy 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyValAlaPheValGluPheAsp 360
Db 1241 AATACATATTGCTCTTTAAGAAATGCTGTAGTATGTTGAGCCCTTTGTAGAAATTTGAC 1300
Qy 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
Db 1301 GTACACCTTTCAAGACATTTGTGCCCGTGGTATATCATCATCTTACCTGTTGTTGACT 1360
Qy 381 MetLysLysValPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
Db 1361 ATGAAAAGAAATTTGATGCTGATCCAGTTGAATTTATTTGAAATTCAGTAAAGAAATTA 1420
Qy 401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
Db 1421 ACATTTGACCACTCCAGTTGTTAAGCTCACATCATGTGACTGCATCGAAATCTAAGAT 1480
Qy 421 ArgLysGluSerValValGlnGluAenSerPheSerGluAenGlnProPheProSer 440
Db 1481 CGGAAGAATCTGTGTTTCAGGAGGAAATTCCTTTTCAGAAATCAGCCATTTCCCTTCT 1540
Qy 441 LeuLysMetValLeuLysSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
Db 1541 CTTAAGATGGTTTAGAGTCTTTGCCAGAGATGTAGGGTTTAACATTTGAAATAAATGG 1600
Qy 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
Db 1601 ATCTGCCAGCAAGGATGGATGTGGATGTGAATCTTATCAACATATTTTGCATGAT 1660
Qy 481 LeuPheLeuAspIleLeuLysThrValLeuGluAenSerGlyLysArgIleVal 500
Db 1661 CTGTTTTTGGATATAATTTTAAAACTGTTTTAGAAAAATCTGGGAAGAGGAGAAATAGTG 1720
Qy 501 PheSerSerPheAspAlaAspLysCysThrMetValArgGlnLysGlnAsnLysTyrPro 520
Db 1721 TTTTCTTCATTTGATGCATATTTGCACAATGGTTCCGCAAAAGCAGAACAAATATCCG 1780
Qy 521 IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer 540
Db 1781 ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTTGAATCTCATGACCTCAGATCT 1840
Qy 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
Db 1841 CGGACAAACCCCATTCATGCAATGAGCTTTGCACAGTTTGAATACTTACTGGGGATAAATGTA 1900
Qy 561 HisThrGluAspLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
Db 1901 CATACTGAAGACTTGCTCAGAAACCCATCTATATTTCAAGAGGCAAAAGCTAAGGGACTA 1960
Qy 581 ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
Db 1961 GTCATATCTCTGGGGTGATGATACCAATGATCCATGATGATGATGATGATGATGATGATGAT 2020
Qy 601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
Db 2021 CTTGGAGTTAATGGTCTAATTTATGATAGATATATGATGATGATGATGATGATGATGATGAT 2080
Qy 621 IlePheGlnValGluGlnLeuGluArgLysGlnGluLeuProGluLeuLysSerCys 640
Db 2081 ATATTCAGAGTGGAGCAATTTGGAACCCCTGAAGCAGGAAATTCGACAGCTTAAGAGCTGT 2140
Qy 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 2141 TTGTGTCCCACTGTTAGCCGCTTTGTTCCTCTATCTTTGTGTGGGAGTCTGATATCCAT 2200
Qy 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2201 GTGGATGCCACCGCATTCATAACGTGGAGATGCT 2236
; Sequence 29632, Application US/10357930
```

RESULT 7

US-10-357-930-29632

; Sequence 29632, Application US/10357930

```
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29632
; LENGTH: 3499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 3490..3491, 3492, 3493, 3494, 3495, 3496, 3497, 3498, 3499
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29632
Alignment Scores:
Pred. No.: 0 Length: 3499
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0
US-10-047-855-3 (1-672) x US-10-357-930-29632 (1-3499)
Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 221 ATGACACCTTCTCAGGTTGCCTTTGAAATAAGAGGAACCTCTTTTACGAGGAGAAGTTTTT 280
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 281 CGGATATGGAAGCTGTGATGCTTTGGGAAATCGGAATCCTCAAAATGCTGGGCTCTT 340
Qy 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 341 CTTCCAGAGAATGACACAGCTGAAGCATGCTATGGAAAGCAACCATTTGACTCACTAGA 400
Qy 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 401 CGGATATCAGTTTCAGTATCGCTACTTCAAGGGGTACTTTTTTGAACCAAGACTATCGGT 460
Qy 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 461 GTTCCATGTCAAGTGATAGTTTCAAGTGGGAGACTCATCTCAACCCAGATCAATAACC 520
Qy 101 ProLeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisGlnGlyValGlu 120
Db 521 CTTTGAAGGCAATTAATTATTATGCGATGACCAATTTGGAATCCCAATGGTGTGAA 580
Qy 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
Db 121 |||||
```


581 ACTCGGATTCGTGATGGCTGACATGTCAGACTGAAATAGATTACGTTTGCATTATCT 640 Db
141 GluLysProValSerIleThrLysLysLysLysLysLysSerArgPheArgValLys 160 Qy
641 GAAAAACCTCTGTGTCATTAACCAAGAAAAATTAATAAATCTAGATTAGGGTGAAG 700 Db
161 LeuThrLeuGluGlyLeuGluGluAspAspArgValSerProThrValLeuHis 180 Qy
701 CTGACACTAGAAAGGCTGGAGGAAGATGACGATGATAGGGTATCTCCACCTGTACTCCAC 760 Db
181 LysMetSerAsnSerLeuGluLysSerLeuLysSerAspAsnGluPheLysCysArgHis 200 Qy
761 AAAATGTCCTAGCTGGAGATATCTTAATAAGCGCAATGAGTTCAGTGGAGGAT 820 Db
201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTTPThrGluTyrSerIleGln 220 Qy
821 TCACAGCCGGAGTGTGGTATGGCTTGAGCTGATCGTTGGACAGAGTACAGCATACAG 880 Db
221 ThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGlu 240 Qy
881 ACGATGGAACACAGATAACCTGGAATCTTTGATTTTTTCGAAGAAGATCTCAGTGAG 940 Db
241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260 Qy
941 CACGTAGTTTCAGGGTATGCCCTTCTGGACATGGGTACAGCTTGTCTCTTATCATCC 1000 Db
261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280 Qy
1001 ACATTCCTGAGAGTGGAAAGAGTGTGGAATCTTACTCTTCCCATCATGAGCAGAAAT 1060 Db
281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300 Qy
1061 TCCCGGAAAAACAATAGGCAAGTGAGAGTTGACTATATAATTAATTAAGCCATTTACCGGA 1120 Db
301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320 Qy
1121 TACAGTTGTGACATGAATCTTCAATTTCCAAAGTATTGGAAGCCAAAGATACCATTTGAT 1180 Db
321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340 Qy
1181 GTTGGCCATCGAGTGCAGGAACTCTACACACTGCCAGCTGGCTAAAGTTCAAGAA 1240 Db
341 AsnThrIleAlaSerLeuArgAsnAlaLaserHisGlyAlaAlaPheValGluPheAsp 360 Qy
1241 AATACTATTGCTTCTTTAAGAAATGCTGCTAGTGTGAGTGCAGCCCTTTGTAGAATTGAC 1300 Db
361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380 Qy
1301 GTACACCTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTGTTGACT 1360 Db
381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400 Qy
1361 ATGAAAAAGAAATTTGATGCTGATCCAGTTGAATATTGAAATTCAGTAAAGAAATTA 1420 Db
401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420 Qy
1421 ACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGCTGACTGCACTGAATCTAAGGAT 1480 Db
421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440 Qy
1481 CGAAAAAGAAATCTGTGGTTCAGGAGGAAAAATTCCTTTTCAGAAATACGCCATTTCTTCT 1540 Db
441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460 Qy
1541 CTTAAGATGGTTTTAGAGTCTTTGCCAGAGAGTGTAGGGTTTAACTTGAATTAATAATGG 1600 Db
461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrThrPheAspMetAsn 480 Qy
1601 ATCTGCCAGCAAGAGGATGGAATGTGGATGGTAACTTATCAACATATTTTGACATGAAT 1660 Db
481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal 500 Qy
1661 CTGTTTTTGGATATAATTTTAAAAAATCTGTTTTAGAAAAATCTGGGAAGAGGAGATAGTG 1720 Db

501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro 520 Qy
1721 TTTTCTTCATTGATGTCAGATATTTGCAATAGTTTGGCAAAAGCAGAAACAAATATCG 1780 Db
521 IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer 540 Qy
1781 ATACTATTTTAACTCAAGGAAAACTGAGATTTATCTCTGAACTCATGACCTCAGATCT 1840 Db
541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560 Qy
1841 CGGACAAACCCCATTCGAATGAGCTTTGACAGTTTGAANAATCTACTGGGATAAATGTA 1900 Db
561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580 Qy
1901 CATACTGAAGACTTGTTCAGAAACCCCATCTATATTCAAGAGGCAAAAGCTTAAGGACTA 1960 Db
581 ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600 Qy
1961 GTCATATTTCTGCTGGGTGATGATACCAATGATCTCTGAAACAGAAAGGAAATTAAGGAA 2020 Db
601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620 Qy
2021 CTTGGAGTTAATGGTCTAAATTTATGATAGATATATGATTTGGATGCTGACACCAAT 2080 Db
621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnLeuLeuProGluLeuLysSerCys 640 Qy
2081 ATATTCAGTGGAGCAATTTGGAACGCTGAGCAGGAAATTCAGAGCTTAAGAGCTGT 2140 Db
641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660 Qy
2141 TTGTGTCCCACTGTTAGCCGCTTGTTCCTCATCTTTGTGTGGGAGTCTGATATCCAT 2200 Db
661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672 Qy
2201 GTGGATGCCACGCGCATTTGATTAACGTGGAGAAATGCT 2236 Db

RESULT 8

US-10-426-776-39
; Sequence 39, Application US/10426776
; Publication No. US20040009553A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Williams, Mark J.
; APPLICANT: Tsia, Pong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Chiang, Lillian Wei-Ming
; APPLICANT: Hunter, John Joseph
; APPLICANT: Wood, Andrew
; APPLICANT: Jenkins, Lorayne P.
; TITLE OF INVENTION: NOVEL 27411, 23413, 22438, 23553,
; TITLE OF INVENTION: 25278, 26212, NARC SCI, NARC 10A, NARC 1, NARC 12, NARC 13,
; TITLE OF INVENTION: NARC17, NARC 25, NARC 3, NARC 4, NARC 7, NARC 8, NARC 11,
; TITLE OF INVENTION: NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20, NARC 26, NARC
; TITLE OF INVENTION: 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C, NARC
; TITLE OF INVENTION: 8B, NARC 9, NARC2A, NARC 16B, NARC 1C, NARC 1A, NARC 25,
; FILE REFERENCE: 86604 AND 32222 MOLECULES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/10/426,776
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 10/229,662
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/795,691
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,517
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 10/105,992
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/406,045
; PRIOR FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 10/314,881

```

/ PRIOR FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: 09/773,426
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 09/495,823
/ PRIOR FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: 09/692,785
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/161,188
/ PRIOR FILING DATE: 1999-10-22
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 39
/ LENGTH: 3381
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
US-10-426-776-39

Alignment Scores:
Pred. No.: 0 Length: 3381
Score: 2965.50 Matches: 571
Percent Similarity: 90.76% Conservative: 38
Best Local Similarity: 85.10% Mismatches: 33
Query Match: 83.84% Indels: 29
DB: 17 Gaps: 6

US-10-047-855-3 (1-672) x US-10-426-776-39 (1-3381)

QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
DB 186 ATGACACCTTCTCAGGTCACTTTTGAATTAAGAGGAACCTTTTACCAGGAGGCTCTT 245
QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
DB 246 GCATGTGTGGAACTGTGATGCTTGGGAACTGAGTCTCTCAAAATGCTGTGCCT--- 302
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
DB 303 CTTACTGAGATGAGACAGCGGAAGT---GTATGGAAAGCAGTGTGTTCTTAGTAGA 359
QY 61 GlyValSerValGlnTrpArgGlyPheLysGlyTyrPheLeuGluProLysThrIleGly 80
DB 360 GGAATGCTCGTGAAGTACCGCTACTTCAGAGGCTGCTTTTGAACCAAGACTATCGGT 419
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
DB 420 GGTCCATGTCAAGTCACTAGTTCACAGTGGAGACTCACTACCAACCATCAATACC 479
QY 101 ProLeuGluSerGluIleIleLeuAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
DB 480 CCTTTAGAAACCAATCATTTATTGACGATGGACAATTTGGAATCCACAATGCTGTGAA 539
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
DB 540 ACACCTGGATCTCGATGGCTTACCTGTCAAGTGAATTAAGACTGCGTCTGCAATTTTCT 599
QY 141 GluLysProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
DB 600 GAGAAACCTCTCTGTTCAATACCAAGAAAGATTCCAAAATCTAGATTAGGGTAAG 659
QY 161 LeuThrLeuGluGlyLeuGluGlu-----AspAspAspArgValSerProThrVal 178
DB 660 CTTTACACTAGAGGCTCTGGGGAAGATGATGACGACGATGATGAAGCATCTCCCACTGTT 719
QY 179 LeuHisLysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCys 198
DB 720 CTTTCAAGATGTCCCAATAGCTGGAGATATCTTAATTAAGTCAACATCAAGTTCGAAGTGC 779
QY 199 ArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSer 218
DB 780 AGGCACCTCAGCCAGAAATGTGGGTATGGCTTACAGCTGACCGCTGGACAGATACAGC 839
QY 219 IleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeu 238

```

```

DB 840 ATACAGACAATGAGCCGACCAACCTTGAACCTCATCTTTGACTTTTGGAGAGATCTC 899
QY 239 SerGluHisValValGlnGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeu 258
DB 900 AGTGACCATGTAGTCCAGGGTGATGTTCTTCTGGACATGTGGGACAGCATGCTCTCTG 959
QY 259 SerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProfileMetSer 278
DB 960 TCATCTACCATTGCTCAGAGTGAAGAGCGCTGGAATCCTTACTCTTCCCATCATGAGC 1019
QY 279 ArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeu 298
DB 1020 AGAAGTTCAGAAAACTATAGCAAAAGTCAGAGTTGATTTATATCATCATCAAGCCATTA 1079
QY 299 ProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIlePro 318
DB 1080 CAGGATATAGTTGTTCTATGAGTCTTCACTTCCCAAGTATTTGAACCAACCAATACCA 1139
QY 319 LeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysVal 338
DB 1140 CTGGATGTTGGACATCGTGTGAGGAACTCAACAACAACCTGCCAAGCTGGCTAAAGTA 1199
QY 339 GlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGlu 358
DB 1200 CAGGAAAAATACTATTGCTTCTTTAAGAAATGCTGCCAGCCATGCTGTCAGCATTTGTGAA 1259
QY 359 PheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCys 378
DB 1260 TTTGATGTCACCTTTCAAAGGACTTAGTGCTGTAGTGTATCATCATCTCACCTGCTGT 1319
QY 379 LeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLys 398
DB 1320 TTAACCTATCAAAAGAAATATGAAGCTGATCCAGTTGAATTTGTTGAAATCCCAAGTAAAG 1379
QY 399 GluLeuThrPheAspGlnLeuGlnLeuLysLeuLeuHisValThrAlaLeuLysSer 418
DB 1380 GAATTAACATTCGACCAACTCCAGTATTATTGAAGCTTCTCATGTGATGTCACATAAAC 1439
QY 419 LysAspArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPhe 438
DB 1440 AAAGACCAAAACAATGTATGCTGAGGAGAAATTCCTTTCTGAAAACCAACCATTT 1499
QY 439 ProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIle 458
DB 1500 CCTTCTCTTAAGATGCTTTTAGAGTCAATGCCAGAAAAATGTAGGATTTAATATAGAAATA 1559
QY 459 LysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAsp 478
DB 1560 AAATGGATTTGCCAACACACAGGATGAGTATGGACGCAACTTATCGACATATTTGAT 1619
QY 479 MetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArg 498
DB 1620 ATGAATGCAATTTTGGATATAATTTTAAAAACTGTTTAAAGAAATTCGCGGAAAGAGAGA 1679
QY 499 IleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLys 518
DB 1680 ATAGTATTTTCTTCAATTCATTCAGACATCTGTACAAATGTTTGGCAGAAACAAACAAA 1739
QY 519 TyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeu 538
DB 1740 TATCCCATATATTTTGAACCAAGGAAAGTCTGACATTTTACCCTGAACCTCATGGACCTC 1799
QY 539 ArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluLeuLeuGlyIle 558
DB 1800 AGATCTCGACAAACCCCATTCGAATGAGCTTTGACAGTTCGAAATATATTTTGGGGATA 1859
QY 559 AsnValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLys 578
DB 1860 AATGCCCATCTAGAGATCTCTTAGAAACCCCATCTATGCTCCAGAGGCAAGATAG 1919
QY 579 GlyLeuValIlePheCysTrpGlyAspThrAsnAspProGluAsnArgArgLysLeu 598

```

```

Db 1920 GGATGGTCATATCTGCTGGGGTGATGATACCAATGATCTGAAAAACAGAAACTG 1979
Qy LysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGln 618
Db 1980 AAGGAATTGGAGTAATGGCTTAATATATATAGTAT----- 2018
Qy 619 ProAsnIlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeu----- 634
Db 2019 -----TTGTTTGTGTAATAAATCTCCATGGAATTTGTTCAACAGTGTAGTTTATCTA 2072
Qy 635 -----ProGluLeuLysSer-----CysLeuCys 642
Db 2073 TTTTAACTATTTTAAATATAGATAGTTAGCTTAAAGTTTATCTTGACACTGTGACCTT 2132
Qy 643 ProThrValSerArgPheValProSerSerLeu 653
Db 2133 TCCAGGTGTGAGATATGTCAAAAGCCACTTA 2165

RESULT 9
US-10-426-776-32
; Sequence 32, Application US/10426776
; Publication NO. US2004000953A1
; GENERAL INFORMATION:
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Williams, Mark J.
; APPLICANT: Tsia, Fong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Chiang, Lillian Wei-Ming
; APPLICANT: Hunter, John Joseph
; APPLICANT: Wood, Andrew
; APPLICANT: Jenkins, Lorayne P.
; TITLE OF INVENTION: NOVEL 27411, 23413, 22438, 23553,
; TITLE OF INVENTION: 25278, 26212, NARC SCI, NARC 10A, NARC 1, NARC 12, NARC 13,
; TITLE OF INVENTION: NARC17, NARC 25, NARC 3, NARC 4, NARC 7, NARC 8, NARC 11,
; TITLE OF INVENTION: NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20, NARC 26, NARC
; TITLE OF INVENTION: 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C, NARC
; TITLE OF INVENTION: 8B, NARC 9, NARC2A, NARC 16B, NARC 1C, NARC 1A, NARC 25,
; TITLE OF INVENTION: 86604 AND 32222 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0620NMIM
; CURRENT APPLICATION NUMBER: US/10/426, 776
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 10/229, 662
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/795, 691
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185, 517
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 10/105, 992
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/406, 045
; PRIOR FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 10/314, 881
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 09/773, 426
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 09/495, 823
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/692, 785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/161, 188
; PRIOR FILING DATE: 1999-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 2738
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-426-776-32
Alignment Scores:

```

```

Pred. No.: 1.39e-288 Length: 2738
Score: 2552.00 Matches: 486
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.15% Indels: 0
DB: 17 Gaps: 0

US-10-047-855-3 (1-672) x US-10-426-776-32 (1-2738)
Qy 187 GluIleSerLeuIleSerAspGlnPheLysCysArgHisSerGlnProGluCysGly 206
Db 18 GAGATATCTTAAATAAGCACAATGAGTTCAGTGCAGGCANTTACAGCGGAGTGTGCT 77
Qy 207 TyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGlnThrMetGluProAspAsn 226
Db 78 TATGGCTTTCAGCGCTGATGTTGGACAGAGTACAGCATACAGACGATGAAACCCAGATAAC 137
Qy 227 LeuGluLeuIlePheAspPhePheGluGluAspLeuSerGluHisValGlnGlnGlyAsp 246
Db 138 CTGGAACCTAATCTTTGATTTTTTCGAAAGAAGATCTCAGTGAGCACGTAGTTCAGGGTGAT 197
Qy 247 AlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAlaGluSerGly 266
Db 198 GCCCTTCCTGGACATGCGGTACAGCTTGTCTTATCATCCACCATTTGCTGAGAGTGA 257
Qy 267 LysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsnSerArgLysThrIleGly 286
Db 258 AAGAGTCTGGAATTCCTACTCTCCATCATGACGAGAATTTCCCGGAAAAACAATAGGC 317
Qy 287 LysValArgValAspTyrIleIleLysProLeuProGlyTyrSerCysAspMetLys 306
Db 318 AAGTGAGAGTTGACTATATAATTAATTAAGCCATTACAGATACAGTTCGTGACATGAA 377
Qy 307 SerSerPheSerLysTyrTrpLysProArgIleProLeuAspValGlyHisArgGlyVala 326
Db 378 TCTTCATTTTCCAGTATTGGAAGCCAAAGTAATCCATGGATGTTGGCCATCGAGGTGCA 437
Qy 327 GlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeu 346
Db 438 GGAACCTCTACACACTGCCAGCTGGCTAAAGTTCAAGAAAAATACTATTGCTTCTTTA 497
Qy 347 ArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAsp 366
Db 498 AGAAATGCTGCTAGTGTGACGCTTTGTAGAAATTTGACGTACACCTTTCAAGAGAC 557
Qy 367 PheValProValValTyrHisAspLeuThrCysCysValLeuThrMetLysLysLysPheAsp 386
Db 558 TTTGTGCCCCGGGTATATCATGATCTTACCTGTTGTTGACTATGAAAAAGAAATTTGAT 617
Qy 387 AlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAspGlnLeuGln 406
Db 618 GCTGATCCAGTTGAATTAATTTGAAATTCAGTAAAAAGAAATTAACATTTGACCACTCCAG 677
Qy 407 LeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSerValVal 426
Db 678 TTGTTAAAGCTCAGTCATGCTGACTGCACTGAAATCTAAGGATCGGAAAGAAATCTGTGTT 737
Qy 427 GlnGluGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMetValLeuGlu 446
Db 738 CAGGAGGAAAAATTCCTTTTCAGAAAAATCAGCCATTTCTTCTTTAAGATGTTTTAGAG 797
Qy 447 SerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGlnGlnArgAsp 466
Db 798 TCTTTGCCAAGAGTGTAGGGTTTAAACATTGAAATAAAAATGATCTGCGCAGCAAAAGGAT 857
Qy 467 GlyMetTrpAspGlyAsnLeuSerThrThrPheAspMetAsnLeuPheLeuAspIleIle 486
Db 858 GGAATGCTGGGATGGTAACTTATCAACATATTTTGACATGAATCTGTCTTTTGTATATAAT 917
Qy 487 LeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSerPheAspAla 506
Db 918 TTAATAAACTGTTTTAGAAAAATTTCTGGGAAAGAGAGATAGTGTCTTTCTTCTTGTATGCA 977

```

```
QY 507 AspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeuThrGln 526
Db 978 GATATTGACAAATGGTTCGGCAAAAGCAGAAACAAATATCCGACTACTATTTTAACTCAA 1037
QY 527 GlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThrProIleAla 546
Db 1038 GGAAATCTGAGATTATCTGAACTCATGGACCTCAGATCTCGGACAAACCCCATGCA 1097
QY 547 MetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGluAspLeuLeu 566
Db 1098 ATGAGCTTTGCACAGTTTGAAATCTACTGGGATAAATGTACATACTGAAGACTTGCTC 1157
QY 567 ArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpGly 586
Db 1158 AGAAACCCATCTCTATATCAAGAGGCAAAAGCTTAAGGGACTAGTCATATTTCTCTGGGT 1217
QY 587 AspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyValAsnGlyLeu 606
Db 1218 GATGATACCAATGATCTGAAACAGAGAAATTTGAGAACTTGAGTTAATGTCTA 1277
QY 607 IleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePheGlnValGluGln 626
Db 1278 ATTATGATAGATATATGATTGGATGCTGACACCAACCAATATATTCAGTGGAGCAA 1337
QY 627 LeuGluArgLysGlnGluLeuProGluLeuLysSerCysLeuCysProThrValSer 646
Db 1338 TTGGAACGCTGAGCAGCAAAATGCGCAGAGCTTAAGAGCTGTTGTGTCCTGCTTAGC 1397
QY 647 ArgPheValProSerSerLeuCysGlyGluSerAspIleHisValAspAlaAsnGlyIle 666
Db 1398 CGCTTGTTCCTTCATCTTGTGTGGGAGTCTGATATCATGTGGATGCCAACGCAATT 1457
QY 667 AspAsnValGluAsnAla 672
Db 1458 GATAAGCTGAGATGCT 1475
RESULT 10
US-10-426-776-48
; Sequence 48, Application US/10426776
; Publication No. US20040009553A1
; GENERAL INFORMATION:
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Williamson, Mark J.
; APPLICANT: Tsia, Fong-Ying
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Chiang, Lillian Wei-Ming
; APPLICANT: Hunter, John Joseph
; APPLICANT: Wood, Andrew
; APPLICANT: Jenkins, Lorayne P.
; TITLE OF INVENTION: NOVEL 27411, 23413, 22438, 23553,
; TITLE OF INVENTION: 25278, 26212, NARC SCI, NARC 10A, NARC 1, NARC 12, NARC 13,
; TITLE OF INVENTION: NARC17, NARC 25, NARC 3, NARC 4, NARC 7, NARC 8, NARC 11,
; TITLE OF INVENTION: NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20, NARC 26, NARC
; TITLE OF INVENTION: 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C, NARC
; TITLE OF INVENTION: 8B, NARC 9, NARC2A, NARC 16B, NARC 1C, NARC 1A, NARC 25,
; TITLE OF INVENTION: 86604 AND 32222 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MP103-0620NIM
; CURRENT APPLICATION NUMBER: US/10/426, 776
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 10/229,662
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/795,691
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,517
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 10/105,992
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/406,045
; PRIOR FILING DATE: 1999-09-27
; PRIOR APPLICATION NUMBER: 10/314,881
; PRIOR FILING DATE: 2002-12-09
```

```
; PRIOR APPLICATION NUMBER: 09/773,426
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 09/495,823
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 09/692,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/161,188
; PRIOR FILING DATE: 1999-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 2393
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-426-776-48
Alignment Scores:
Pred. No.: 4,73e-153 Length: 2393
Score: 1399.00 Matches: 261
Percent Similarity: 95.85% Conservative: 16
Best Local Similarity: 90.31% Mismatches: 12
Query Match: 39.55% Indels: 0
DB: 17 Gaps: 0
US-10-047-855-3 (1-672) x US-10-426-776-48 (1-2393)
QY 384 LysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAsp 403
Db 71 AAATATGAAGCTGATCCAGTTGATTTTGAATCCAGTAAGAGAAATTAACATTCGAC 130
QY 404 GlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGlu 423
Db 131 CAATCCAGTTATTGAAGCTTCTCATGTGACTGTCACATAAAACCAAGACCAAGAACAA 190
QY 424 SerValValGlnGluLysSerPheSerGluAsnGlnProPheProSerLeuLysMet 443
Db 191 TGTATGGCTGAGAGGAGAAATTCCTTTCTGAAACCAACCATTTCTCTCTTAAGATG 250
QY 444 ValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGln 463
Db 251 GTTTTAGAGTCATTTGCCAGAAATTTAGAGATTATATAGAAATAAATATGGATTCGCAA 310
QY 464 GlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeu 483
Db 311 CACAGGATGGAGTATGGAGCGCAACTTATCGACATATTTTGATATGAATGCAATTTTG 370
QY 484 AspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSer 503
Db 371 GATATAATTTTAAAAACTGTTTTAGAAAATTCGGGAGAGAGAGATAGTATTTTCTTCA 430
QY 504 PheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPhe 523
Db 431 TTGTATGCAGACATCTGTACAAATGGTTCGGCAGAAACAAACAAATATCCCATATTATT 490
QY 524 LeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThr 543
Db 491 TTGACCCCAAGGAAGCTGCACATTTACCTTGAACTCATGCAGCTCAGATCTCGGACACA 550
QY 544 ProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGlu 563
Db 551 CCCATTGCAATGAGCTTTGCACAGTTTGAAAAATATTTTGGGATTAATGCCATCTGAA 610
QY 564 AspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyValIlePhe 583
Db 611 GATCTCTTAGAAACCCATCTTATGTCCAAGAGGCAAAAGATAAGGGATTTGGTCATATC 670
QY 584 CysTrpGlyAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyVal 603
Db 671 TGCTGGGTGATGATACCAATGATCTGAAACAGAGAGAACTGAAGGAATTTGGAGTA 730
QY 604 AsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePheGln 623
```

Db 731 AATGGCTTAATATATAGATAGATATAGATGGATGCTGAACACCAATATATATCCAA 790
 Qy 624 ValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCysLeuCysPro 643
 Db 791 GTGAGCACTGGAGCGCTGAAGCGAGAAATTCAGAGCTTAAGAACTGTTGTGTCCC 850
 Qy 644 ThrValSerArgPheValProSerSerLeuCysGlyGluSerAspLleHisValAspAla 663
 Db 851 ACTGTTAGCCACTTCATCTCTCTCTGTATGGAGTCTAAAAATCCATGTGGATGCT 910
 Qy 664 AsnGlyLleAspAsnValGluAsnAla 672
 Db 911 AACGCATTGATAATGTGGAGAACGCT 937
 RESULT 11
 US-09-998-598-1849
 ; Sequence 1849, Application US/09998598
 ; Patent No. US20020150922A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Chenault, Ruth A.
 ; APPLICANT: Meagher, Madelein Joy
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
 ; FILE REFERENCE: 210121.561
 ; CURRENT APPLICATION NUMBER: US/09/998,598
 ; CURRENT FILING DATE: 2001-11-16
 ; NUMBER OF SEQ ID NOS: 2606
 ; SOFTWARE: Corixa Invention Disclosure Database
 ; SEQ ID NO 1849
 ; LENGTH: 471
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-998-598-1849
 Alignment Scores:
 Pred. No.: 1.4e-86 Length: 471
 Score: 824.00 Matches: 156
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 23.30% Indels: 0
 DB: 9 Gaps: 0
 US-10-047-855-3 (1-672) x US-09-998-598-1849 (1-471)
 Qy 166 LeuGluClnAspAspAspArgValSerProThrValLeuHisLysMetSerAsnSer 185
 Db 2 CTGGAGGAAGATGACGATGATGGGTATCTCCCACTGTACTCCACAAATGTCCATAGC 61
 Qy 186 LeuGluLleSerLeuLleSerAspAsnGluPheLysCysArgHisSerGlnProGluCys 205
 Db 62 TTGAGATATCTTAAATAGCGACATGAGTTTAAGTCGACGATTCACAGCGGAGTGT 121
 Qy 206 GlyTyrGlyLeuGlnProAspArgTfThrGluTyrSerIleGlnThrMetGluProAsp 225
 Db 122 GGTATGCTTGACGCTCGTTCGTGACAGAGTACAGCATACAGACGATGGAACAGAT 181
 Qy 226 AsnLeuGluLlePheAspPhePheGluAspLeuSerGluHisValValGlnGly 245
 Db 182 AACCTGGAACCTAATCTTTGATTTTTCGAAGAAGATCTCAGTAGCACGTAGTTCCAGGT 241
 Qy 246 AspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAlaGluSer 265
 Db 242 GATGCCCTTCTCGACATGCGGTACAGCTGTCTCTTAATCATCCACCATTCGTGAGAGT 301
 Qy 266 GlyLysSerAlaGlyLleLeuThrLeuProLleMetSerArgAsnSerArgLysThrIle 285
 Db 302 GGAAAGAGTGTGGAATCTTACTCTTCCATCATGACGAGAAATCCCGGAAACAAATA 361
 Qy 286 GlyLysValArgValAspTyrIleIleLysProLeuProGlyTyrSerCysAspMet 305
 Db 362 GCGAAAGTGAGAGTTGACTATATATTAATTAAGCCATTCACGAGATACAGTTGTGACATG 421

Qy 306 LysSerSerPheSerLysTyrTrpLysProArgLleProLeuAspVal 321
 Db 422 AAATCTTCATTTTCCAGTATTGGAAGCCCAAGAAATACCATTCGATGTT 469
 RESULT 12
 US-09-814-353-16669
 ; Sequence 16669, Application US/09814353
 ; Publication No. US20030165831A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lee, John
 ; APPLICANT: Thompson, Pamela
 ; APPLICANT: Lillie, James
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
 ; FILE REFERENCE: MRI-006B
 ; CURRENT APPLICATION NUMBER: US/09/814,353
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/191,031
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: US 60/207,124
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: US 60/211,940
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: US 60/216,820
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: US 60/220,661
 ; PRIOR FILING DATE: 2000-07-25
 ; PRIOR APPLICATION NUMBER: US 60/257,672
 ; PRIOR FILING DATE: 2000-12-21
 ; NUMBER OF SEQ ID NOS: 22037
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16669
 ; LENGTH: 520
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-814-353-16669
 Alignment Scores:
 Pred. No.: 6.34e-74 Length: 520
 Score: 717.00 Matches: 137
 Percent Similarity: 99.28% Conservative: 1
 Best Local Similarity: 98.56% Mismatches: 1
 Query Match: 20.27% Indels: 0
 DB: 10 Gaps: 0
 US-10-047-855-3 (1-672) x US-09-814-353-16669 (1-520)
 Qy 371 ValTyrHisAspLeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspProVal 390
 Db 104 GTATATCATGATCTTACCTGTTGTTGACTATGAAAGAAATTTGATGCTGATCCAGTT 163
 Qy 391 GluLeuPheGluLleProValLysGluLeuThrPheAspGlnLeuGlnLeuLysLeu 410
 Db 164 GAATTTATTGAAATTCAGTAAAGAAATTAACATTTGACCACTCCAGTTGTTAAAGCTC 223
 Qy 411 ThrHisValThrAlaLeuLysSerLysAspArgLysGluSerValValGlnGluGln 430
 Db 224 ACTCATGTGACTGCACCTGAAATCTAAGGATCGAAAGAAATCTGTGTTGTCAGGAGGAAAT 283
 Qy 431 SerPheSerGluAsnGlnProPheProSerLeuLysMetValLeuGluSerLeuProGlu 450
 Db 284 TCTTTTCGAAAGAAATCAGCCATTTCTCTCTTTAAGATGGTTTTAGAGTCTTTGCCAGAA 343
 Qy 451 AspValGlyPheAsnLleGluLleLysTrpLleCysGlnArgAspGlyMetTrpAsp 470
 Db 344 GATGTAGGGTTTAAACATTCGAAATGGAATCTGCCAGCAAGGATGGAATGCGGAT 403
 Qy 471 GlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeuAspIleLleLeuLysThrVal 490
 Db 404 GGTAACCTTATCAACATATTTTGACATGAATCTGTTTTCGATATATATTTTAAAAAAGTGT 463

QY 491 LeuGluAenSerGlyLysArgIleValPheSerSerPheAspAlaAspIleCys 509
Db 464 TTAGAAATTCGGGAAGAAGAAATAGTGGTTCTTCATTGATGAGATATTTGC 520
RESULT 13
US-09-796-692-8067
; Sequence 8067, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8067
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (11)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (16)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (184)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (188)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (213)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (298)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (346)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8067
Alignment Scores: 8.14e-72 Length: 436
Pred. No.:

Score: 698.00 Matches: 138
Percent Similarity: 94.52% Conservative: 0
Best Local Similarity: 94.52% Mismatches: 7
Query Match: 19.73% Indels: 1
DB: Gaps: 0
US-10-047-855-3 (1-672) x US-09-796-692-8067 (1-436)
QY 335 LeuAlaLysValGlnGluAenThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAla 354
Db 1 CTGGCTAAAGNTCAANAANAATACTATTGCTTCTTTAAGAAATGCTGCTAGTCATGTCGA 60
QY 355 AlaPheValGluPheAspValHisLeuSerLysAspPheValProValValTyrHisApp 374
Db 61 GCCTTTGTAGAAATTCACGTACACCTTTCAAAGGACTTTGTGCCCGCTGTATATCATGAT 120
QY 375 LeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGlu 394
Db 121 CTTACCTGTGTTGACTATGAAAGAAATTTGATGCTGATACAGTTGAATTTTGA 180
QY 395 IleProValLysGluLeuThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThr 414
Db 181 ATTNCAGNAAGAATAAACAATTCACCACTNCAGTTGTTAAAGCTCACTCATGTGACT 240
QY 415 AlaLeuLysSerLysAspArgLysGluSerValValGlnGluAenSerPheSerGlu 434
Db 241 GCACCTGAAATCTAAGATCGAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTTTCANAA 300
QY 435 AsnGlnProPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPhe 454
Db 301 AATCAGCCATTTCTTCTCTTAAGATGGTTTTAGAGTCTTTGCCANAAAGATGTAGGGTTT 360
QY 455 AsnIleGluIleLysTrpIleCysGlnGlnArgAspGlyMetTrp-AspGlyAsnLeuSe 474
Db 361 AACATTGAAATTAATAATGGATCTGCCAGCAAGGATGGATGTGGGATGGTAACTTATC 420
QY 474 rThrTyrPheAspMet 479
Db 421 AACATATTTGCATG 436
RESULT 14
US-10-040-862-8067
; Sequence 8067, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy,
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14

;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8067
;; LENGTH: 436
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (11)
;; OTHER INFORMATION: n=A,T,C or G
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (16)
;; OTHER INFORMATION: n=A,T,C or G
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (184)
;; OTHER INFORMATION: n=A,T,C or G
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (188)
;; OTHER INFORMATION: n=A,T,C or G
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (213)
;; OTHER INFORMATION: n=A,T,C or G
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (298)
;; OTHER INFORMATION: n=A,T,C or G
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (346)
;; OTHER INFORMATION: n=A,T,C or G

US-10-040-862-8067

Alignment Scores:
Pred. No.: 436
Score: 698.00
Percent Similarity: 94.52%
Best Local Similarity: 94.52%
Query Match: 19.73%
DB: 14

US-10-047-855-3 (1-672) x US-10-040-862-8067 (1-436)

QY 335 LeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAla 354
Db 1 CTGGCTAAAGNCTCAANAAATACTATGCTCTTTAAGAAATGCTGCTAGTCATGCTGCA 60
QY 355 AlaPheValGluPheAspValHisLeuSerLysAspPheValProValValTyrHisAsp 374
Db 61 GCCTTTGTAGAAATTTGACGTACACCTTTCAAAGGACTTTGTGCCCGGGTATATCATGAT 120
QY 375 LeuThrCysCysLeuThrMetLysLysPheAspAlaAspProValGluLeuPheGlu 394
Db 121 CTTACCTGTTGTTGACTATGAAAAGAAATTTGATGCTGATACAGTTGAATTATTGAA 180
QY 395 IleProValLysGluLeuThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThr 414
Db 181 ATTNCAGNAAAGAAATTAACATTTGACCAACTNCAGTTGTTAAAGCTCACTCATGTGACT 240
QY 415 AlaLeuLysSerLysAspArgLysGluSerValValGlnGluGluAsnSerPheSerGlu 434
Db 241 GCACCTGAAATCTAAGGATCGAAGAAATCTGTGGTTCCAGGAGAAATTCCTTTTCANAA 300

QY 435 AsnGlnProPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPhe 454
Db 301 AATCAGCCATTTCCTTCTCTTAAGATGGTTTAGAGTCTTTGCCANAAAGATGTAGGGTTT 360
QY 455 AsnIleGluIleLysTrpIleCysGlnGlnArgAspGlyMetTrp-AspGlyAsnLeuSe 474
Db 361 AACATTGAAATAAATGGATCTGCCAGCAAGGATGGAATGTGGGATGGTAACCTTATC 420
QY 474 rThrTyrPheAspMet 479
Db 421 AACATATTTTGACATG 436

RESULT 15

US-10-057-475B-8067
; Sequence 8067, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8067
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(436)
; OTHER INFORMATION: n = g, a, c or t
; US-10-057-475B-8067

Alignment Scores:
Pred. No.: 8,14e-72
Score: 698.00
Percent Similarity: 94.52%
Best Local Similarity: 94.52%
Query Match: 19.73%
DB: 17
Length: 436
Matches: 138
Conservative: 0
Mismatches: 7
Indels: 1
Gaps: 0

US-10-047-855-3 (1-672) x US-10-057-475B-8067 (1-436)

```
QY 335 LeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaSerHisGlyAla 354
Db |||||
1 CTGGCTAAAGNTCAANAAATACTATTGCTCTTTAAGAAATGCTGCTAGTCATGGTGCA 60
QY 355 AlaPheValGluPheAspValHisLeuSerLysAspPheValProValValTyrHisAsp 374
Db |||||
61 GCCTTTGTAGAAATTTGACGTCACACCTTTCAAGGACTTTGTGCCCGTGGTATATCATGAT 120
QY 375 LeuThrCysCysLeuThrMetLysLysPheAspAlaAspProValGluLeuPheGlu 394
Db |||||
121 CTTACCTGTTGTTGCTATGAAAGAAATTTGATGCTGATACAGTTGAATTTATTGAA 180
QY 395 IleProValLysGluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHisValThr 414
Db |||||
181 ATTCAGNAAAGAAATTAACATTTGACCACTNCAGTTGTTAAAGCTCACTCATGTGACT 240
QY 415 AlaLeuLysSerLysAspArgLysGluSerValValGlnGluAsnSerPheSerGlu 434
Db |||||
241 GCACTGAAATCTAAGGATCGAAAAGAAATCTGTGGTTTCAGGAGGAAAAATTCCTTTTCANAA 300
QY 435 AsnGlnProPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPhe 454
Db |||||
301 AATCAGCCATTTCCCTTCTCTTAAGATGGTTTTAGAGTCTTTGCCANAAGATGAGGGTTT 360
QY 455 AsnIleGluIleLysTrpIleCysGlnGlnArgAspGlyMetTrp-AspGlyAsnLeuSe 474
Db |||||
361 AACATTGAAATAAAATGGATCTGCCAGCAAGGATGGAATGTGGGGATGTTAACTTATC 420
QY 474 rThrTyrPheAspMet 479
Db |||||
421 AACATATTTTGACATG 436
```

Search completed: July 3, 2005, 13:26:08
Job time : 1109 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 3, 2005, 06:49:10 ; Search time 307 Seconds
(without alignments)
3581.687 Million cell updates/sec

Title: US-10-047-855-3

Perfect score: 3537

Sequence: 1 MTPSQVAFEIRGTLLPGEVF.....LCGESDIHVDANGIDNVENA 672

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q/cgn2_1/USFTO_spool/US10047855/runat_01072005_155529_2259/app_query.fasta_1.839
-DB=issued Patents NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2o -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10047855 @CNG 1.1 177 @runat_01072005_155529_2259 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	654	18.5	2786	4	US-09-270-767-11893
C 2	516	14.6	1180	4	US-09-270-767-27552
C 3	422	13.1	2421	4	US-09-248-796A-1858
4	163	4.6	1161	4	US-09-328-352-1388
5	153	4.3	720	4	US-09-328-352-1148
6	145	4.1	14570	4	US-09-902-540-1012
7	136.5	3.9	759	4	US-09-543-681A-2048
8	131.5	3.7	1487	4	US-09-949-016-4403
9	128.5	3.7	1629	3	US-09-247-155-71
10	121.5	3.6	2682	3	US-09-105-697-9
11	126.5	3.6	1173	4	US-09-252-991A-1900
12	124	3.5	762	4	US-09-902-540-4278

Sequence 1707, Ap	124	3.5	771	3	US-09-134-001C-1707
Sequence 1212, Ap	124	3.5	26012	4	US-09-902-540-1212
Sequence 1, Appli	121.5	3.4	3393	3	US-09-104-324B-1
Sequence 1, Appli	121.5	3.4	3393	3	US-09-162-713-1
Sequence 212, App	120.5	3.4	7425	3	US-09-453-702B-212
Sequence 2792, Ap	120	3.4	957	4	US-09-107-532A-2792
Sequence 1377, Ap	119	3.4	879	4	US-09-543-681A-1377
Sequence 1289, Ap	115.5	3.3	1803	4	US-09-134-000C-2289
Sequence 498, App	115.5	3.3	2306	4	US-09-799-451-498
Sequence 305, App	115.5	3.3	7861	4	US-09-774-528-305
Sequence 259, App	114	3.2	591	4	US-09-248-796A-259
Sequence 2554, Ap	114	3.2	771	4	US-09-583-110-2554
Sequence 801, App	114	3.2	1599	4	US-09-107-433-801
Sequence 216, App	114	3.2	2651	3	US-08-961-527-216
Sequence 1448, App	113	3.2	2748	4	US-09-949-016-1448
Sequence 1076, Ap	113	3.2	6481	4	US-09-620-312D-1076
Sequence 705, App	112.5	3.2	3099	4	US-09-583-110-705
Sequence 4030, Ap	112.5	3.2	3442	4	US-09-710-279-4030
Sequence 565, App	112.5	3.2	5078	4	US-09-620-312D-565
Sequence 33, Appli	112	3.2	2748	3	US-09-199-290-33
Sequence 1, Appli	112	3.2	2748	4	US-09-821-616-33
Sequence 2260, Ap	112	3.2	3196	3	US-09-704-449-1
Sequence 4746, Ap	111.5	3.2	783	3	US-09-134-001C-2260
Sequence 12, Appli	111	3.1	2602	3	US-09-351-814-12
Sequence 202, App	110	3.1	744	4	US-09-107-532A-202
Sequence 1, Appli	110	3.1	2201	4	US-09-642-000-1
Sequence 638, App	110	3.1	2727	4	US-09-248-796A-638
Sequence 1, Appli	110	3.1	1830121	4	US-09-557-884-1
Sequence 1, Appli	110	3.1	1830121	4	US-09-643-990A-1
Sequence 3679, Ap	109	3.1	756	4	US-09-489-039A-3679
Sequence 3, Appli	108.5	3.1	2631	1	US-08-717-515-3
Sequence 5, Appli	108.5	3.1	3255	1	US-08-717-515-5

ALIGNMENTS

RESULT 1

US-09-270-767-11893/c
; Sequence 11893, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11893
; LENGTH: 2786
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11893

Alignment Scores:
Pred. No.: 4.14e-68 Length: 2786
Score: 654.00 Matches: 201
Percent Similarity: 47.17% Conservative: 124
Best Local Similarity: 29.17% Mismatches: 230
Query Match: 18.49% Indels: 135
DB: 4 Gaps: 26

US-10-047-855-3 (1-672) x US-09-270-767-11893 (1-2786)

Qy	3	ProSerGlnValAlaPheGluLeuArg-----GlyThrLeuLeuProGlyGluValPhe 20
Db	2629	CCCACTCTCGGGAGTTCAATGTGGCTGGAGGTTCCTGGCCGCCGAGGAGGACTG 2570
Qy	21	AlaLeuGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db	2569	GGGCTTACCGCGGATGTGAAGACCTCGCGGAGTGCGAGCTGTCCAGAGATGTGGCTCTA 2510

QY 41 LeuProGluAenAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
 DB 2509 -----GAATCGCTGGACGAGCTCAACTGGCAGCGCACGGTGGCCCTTCAGTCC 2462
 QY 61 GlyValSerValGlnThrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
 DB 2461 TGGCCCGCAGCTGGAGTACCGC-----TACTTTGTCTAGTGGAGATCTCTCT 2414
 QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
 DB 2413 GCTACAGCAG-----ATCCGCGTGGGAACCCATTTCAAGCCAGGCTCCCTGGGA 2360
 QY 101 Pro----- 101
 DB 2359 CCCTGTACGAGCTACAGTCAGCAGCGAGTTGGACGCTTTCGGTATTACCTCGGCAACTCC 2300
 QY 102 -----LeuGluSerGluIleIleAsp----- 109
 DB 2299 GATCTAAAGCCGACGGTGCACCGCGGTGGCTGGAACCATGAGCCATCTTCGAGCTCAAG 2240
 QY 110 -----AspGlyGlnPheGlyIleHisAenGlyValGluThrLeuAspSerGlyTrp 126
 DB 2239 TTCATGGCGAGAGATGTTCAAGTCCAGCAG-----ATCGAGACCTTTGAC----- 2192
 QY 127 LeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerGluLysProPro----- 144
 DB 2191 -----CCCCAGCAGCTCCAGCTGMAAGATTGTGCCGTGGAGAGACCGCTGGCCCTG 2141
 QY 145 -----ValSerIleThrLysLysLysLysLysSerArgPheArgValLysLeuThrLeu 163
 DB 2140 CATGTGGAGTACTCCAAACAGGAGTACGCCAGAGCCAGCTGGAGCTGACGCTACTTTC 2081
 QY 164 GluGlyLeuGluAspAspArgValSerProThrValLeuHisLysMetSer 183
 DB 2080 ---GGAGTGCCTTACACCAAGGGGACATGTTCATCATCAGCTGGCGTGGAG 2024
 QY 184 AsnSerLeuGluLeuSerLeuLeuSerAspAsnGluPheLysCysArgHisSerGlnPro 203
 DB 2023 AGGATGATGAG-----CAGCACTTCCGCTTG 1997
 QY 204 GluCysGlyTyrGlyLeuGlnProAspArgTyrThrGluTyrSerIleGlnThrMetGlu 223
 DB 1996 GAGTGC-----TACAGATGTCAC----- 1976
 QY 224 ProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGluHisValVal 243
 DB 1975 -----GAATCTCTG----- 1967
 QY 244 GlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAla 263
 DB 1966 -----GGCAGTGCACCTCTGTCTACCTCGGACCTTACCGCTACTCGCG 1934
 QY 264 GluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSer---ArgAsnSerArg 282
 DB 1933 -----GGCAGCGAGGAGTGTCTCACCCTGCCGATTAAAGTGGCCAAAGATGCAGAT 1883
 QY 283 LysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGlyTyrSer 302
 DB 1882 GAGACCTGGCGGCTGAGGCTTCCCTATGTCCGGTGCAGCTTACCGCTACTCGCG 1823
 QY 303 CysAspMetLysSerSerPheSerLysTyrTrpLysProArgIlePro---LeuAspVal 321
 DB 1822 CTAGACTTCAAGAACACCTATGCTCACTTACTGGCCCAAGAGCTGGCCCAACCTGGATGTG 1763
 QY 322 GlyHisArgGlyAlaGlyAsnSerThrThrThrAlaGlnLeuAlaLysValGlnGluAsn 341
 DB 1762 GGTCTATCTGGAATGGCAAGAT---TACATTTCAGACGCTCTCTCGGAAAGGAGGAAC 1706
 QY 342 ThrIleAlaSerLeuArgAsnAlaSerHisGlyAlaAlaPheValGluPheAspVal 361
 DB 1705 ACTATCGCGTCTTCTCGTGGCGCCCATGAACATCAGCAGACATGATCGATGTC 1646
 QY 362 HisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMet 381

DB 1645 CATTGACTGCTGATGGTGTGCTGTGATTTATCAGATTTCCGACTGCGAACTGCTCCG 1586
 QY 382 LysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThr 401
 DB 1585 CTTGGCAAGCAGATCAGCAGCGCCAGACAGCTGGAGTACTGCTGATCAAGACATAAAC 1526
 QY 402 PheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArg 421
 DB 1525 TATGACTGCTCAAAAGGCTGGCATCTTCTGTGATCGCA----- 1484
 QY 422 LysGluSerValValGlnGluLeuAsnSerPheSer-----GluAsnGln 436
 DB 1483 -----GGCAGAGTGGAGGATATCCTCGCACACCGCCAGAGCCAGGATGGAAACCCG 1430
 QY 437 ProPheProSerLeu-LysMetValLeuGluSerLeuProGluAspValGlyPheAsnIle 456
 DB 1429 ATATTCCCAACGCTGGTGGAGTACTGGAGAAAGCTGCCCAAGTCACTGGGCATTGATGT 1370
 QY 456 eGluIleLysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTy 476
 DB 1369 GGAGATTAAAGTGG---CCACAGCGTCCCGAGGCG---GGAGGATCAGAGGCTGAGCAAC 1316
 QY 476 rPheAspMetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLy 496
 DB 1315 AATCGACAAGAACTTCTTCGCCGACAAGGTGATCATCAGTGTATCCAGAGGGCTGTGG 1256
 QY 496 sArgArgIleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysG1 516
 DB 1255 CAGACCCGATAATCTTCCAGCTTCGATGCTGATGTCATGTCACGATGCTGAGTTCAAGCA 1196
 QY 516 nAsnLysTyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMe 536
 DB 1195 GAACTCTTCCAGTGTATGTTCTTCACGCGAGGAGAGACGAGAAGTGGCAGCGCTTCTCT 1136
 QY 536 tAspLeuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLe 556
 DB 1135 GGATCTCGAACACGGACCTTTATTGACGCGTAAACAATGCTCAAGCTTTCGAGCTGGC 1076
 QY 556 uGlyIleAsnValHisThrGluAspLeuLeu---ArgAsnProSer---TyrIleGlnG1 574
 DB 1075 TGGCACAGCTCCACACCGCCGAGGACTTCTCTGGGCGAATAATCATCAGAAATGCTCGGAAA 1016
 QY 574 uAlaLysAlaLysGlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAs 594
 DB 1015 AGCCAGGATCTGGCCAAATAGTGTGATCTGGGCGCAGCATTCGCAACTCCAAGGAGCG 956
 QY 594 nArgArgLysLeuLysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTr 614
 DB 955 GGTGCACTTTCACCCGATCGGGGCTACCGCCACGCTGCTACGATCGCATGTATCTCTT 896
 QY 614 pMetProGlu-----GlnProAsnIlePheGlnValGluGlnLeuGlu---ArgLeuLy 631
 DB 895 CATGCGGAGGCGCAAGCGCGA-AGCCTTCTTCAAAATCACCCGACTGATGCGCGAGTTCG 837
 QY 631 sGlnGluLeuProGluLeuLysSer 639
 DB 836 CGGCGCAGTGGCCGATCAAGAAAT 812

RESULT 2
 US-09-270-767-27552/c
 ; Sequence 27552, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 27552
 ; LENGTH: 1180

; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-27552

Alignment Scores:

Pred. No.: 5,04e-52 Length: 1180
Score: 516.00 Matches: 139
Percent Similarity: 51.72% Conservative: 71
Best Local Similarity: 34.24% Mismatches: 143
Query Match: 14.59% Indels: 53
DB: 4 Gaps: 12

US-10-047-855-3 (1-672) x US-09-270-767-27552 (1-1180)

```
QY 216 GluTyrSerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPheGlu 235
DB 1159 GAGTACTCCAAACAGGAGTACGGCAAGAGCCAGCTGGAGCTG-----CAG 1115
QY 236 GluAspLeuSerGluHisValValGlnGlnGlyAspAlaLeuProGlyHisVal----- 252
DB 1114 CCTACTTTCGGAGTGCCTTACACCAAGGGGACATTTGTTCATCTTTCACATCAGCTGCCG 1055
QY 252 ----- 252
DB 1054 CTGGAGAGGATGATGGAGCAGCACTTCGGCTGGAGTGCTACAGATGTCCAAAGCAACTC 995
QY 253 ---GlyThrAlaCysLeuLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIle 271
DB 994 CTGGCAGTGCCACTCTGGTCACCTCGACCTGACT-----GGCAGCGAGGAGTG 944
QY 272 LeuThrLeuProIleMetSer---ArgAsnSerArgLysThrIleGlyLysValArgVal 290
DB 943 CTCACCTCGCATTAAGTCGGCAAGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 884
QY 291 AspTyrIleIleIleLysProLeuProGlyTyrSerCysAspMetLysSerSerPheSer 310
DB 883 CCTATGTCGGTGCGAGCTTACCGCTACTCGCCGCTAGCTCAAGAACACCTATGCT 824
QY 311 LysTyrTrpLysProArgIlePro---LeuAspValGlyHisArgGlyAlaGlyAsnSer 329
DB 823 CACTACTGCCCCAGAGCTGGCCCAACCTGGATGGTGGTGCATCTGCGAATGCAAGAGT 764
QY 330 ThrThrThrAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAla 349
DB 763 ---TACATTTGCACAGCTCTCTGCGAAAGGGAGAACACTATCGCTCTTCTCTGAGCGC 707
QY 350 AlaSerHisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheValPro 369
DB 706 CATGAACATCAGCAGACATGATCGAGTGGATGTCCATTTGACTGCTGATGTGCTT 647
QY 370 ValValTyrHisAspLeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspPro 389
DB 646 GTGATTTATCAGATTTTCGAGCTGCGAATGCTCCGCTGGCAGCAGATCAGCCGCCCA 587
QY 390 ValGluLeuPheGluIleProValLysGluLeuThrPheAspGlnLeuGlnLeuLys 409
DB 586 GACCAGCTGGAGTACGTGCTGATCAAAAGACATAAATATGATGATGCTCAAAAGGCTGCGC 527
QY 410 LeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSerValGlnGluGlu 429
DB 526 ATCTTCTCTGTGATCGCA-----GGCCAAGTGGGGAGTAT 491
QY 430 AsnSerPheSer-----GluAsnGlnProPheProSerLeu-LysMetVal 444
DB 490 CCTTCGCACACGCCCGCCAGGATGGACACCGCATATTCACAGCTGGTGGAGGT 431
QY 444 lleGluSerLeuProGluAspValGlyPheAsnIleGluIleTyrTrpIleCysGlnG 464
DB 430 ACTGGAGAAGCTGCCCAAGTCACTGGGCAATTGATGGAGATTAAGTGG---CCACAGCG 374
QY 464 nArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeuAs 484
DB 373 TCGCCAGGGGC---GGAGGATCAGAGGCTGAGCAACAAATCGACAAGAACTTCTTCGCCGA 317
```

RESULT 3

US-09-248-796A-1858
; Sequence 1858, Application US/09248796A
; Patent No. 6747137

GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1858
; LENGTH: 2421
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1858

Alignment Scores:

Pred. No.: 6.2e-45 Length: 2421
Score: 462.00 Matches: 148
Percent Similarity: 44.40% Conservative: 86
Best Local Similarity: 28.08% Mismatches: 187
Query Match: 13.06% Indels: 106
DB: 4 Gaps: 19

US-10-047-855-3 (1-672) x US-09-248-796A-1858 (1-2421)

```
QY 144 ProValSerIleThrLysLysLysLysLysSerArgPheArgValLysLeuThrLeu 163
DB 910 CCGTGTGAATTAACAAGTCTCTTAGCTAAGAGT---TTTGCTACCGAATG----- 960
QY 164 GluGlyLeuGluAspAspAspArgValSerProThrValLeuHisLysMetSer 183
DB 961 -----GATACCTCTTATCATTTATCAATTTACTGTGCTCATTAATGACA 1005
QY 184 AsnSer---LeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHisSerGln 202
DB 1006 AATAATCCAGTGAACCGCCCGTGGTGTGTGATTTACCATTTAGAGATTTCCATGGA--- 1062
QY 203 ProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGlnThrMet 222
DB 1063 -----AGTGCCACC 1071
```


[illegible]

```

Db      2101 GAAGCTGTCAGACCGCATGTAT- - - - -GATATTAGCAAGTCTTTGTCAG 2148
QY      545 IleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnValHisThrGluAsp 564
Db      2149 AATGCTGTCAGATTTTCTAAATAAATGGAATTTGTTGGGTATTGTATCCAATGCATTGGCA 2208
QY      565 LeuLeuArgAsnProSerTyrIleGlnGluAlaIysAlaIysGlyLeuValIlePheCys 584
Db      2209 TTAATCAAAACACCTAGATTGGCCCAAGTTGTTAAATCGATGGATTTGTTGTTTACC 2268
QY      585 TrpGlyAspAspThrAsnAspProGluAsnArgArgIysLeuLysGluLeuGlyValAsn 604
Db      2269 TATGCTACTGAGAACACCAACCAACCGAATTTGGCTAAATTCAAATGAGAGCTGGTGTGAT 2328
QY      605 GlyLeuIleTyrAspArgIle 611
Db      2329 GCTGTCATTGTGGATAGTGT 2349

RESULT 4
US-09-328-352-1388
; Sequence 1388, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1388
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1388

Alignment Scores:
Pred. No.: 3,11e-09 Length: 1161
Score: 163.00 Matches: 86
Percent Similarity: 37.90% Conservative: 55
Best Local Similarity: 23.12% Mismatches: 119
Query Match: 4.61% Indels: 112
DB: 4 Gaps: 16

US-10-047-855-3 (1-672) x US-09-328-352-1388 (1-1161)
QY      311 LysTyrTrpLysProArgIleProLeuAspValGlyHisArgGlyAlaGlyAsnSerThr 330
Db      109 GAATATCAACTTCTTAAAT- - -CTAGTAGTAGGACACCGCGCGCTAGC- - - - - 156
QY      331 ThrThrAlaGlnLeuAlaIysValGlnGluAsnThrIleAsnSerLeuArgAsnAlaA 350
Db      157 - - - - -GCTTTACGTCCTGAACATACTTTAGCTTCATATCAAAAGCGATT 201
QY      351 SerHisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheValProVal 370
Db      202 GATGACGCGCAGATTTTCAITGAACCGGATCTAGTCTCTACAAAGATGCGGTACTGGTT 261
QY      371 ValTyrHisAsp- - - - - - - - - - -LeuThrCysCysLeu 379
Db      262 GCCCGCCCATGAATGAATTTGGTGGAAACCAACCAATGTAAGCACTTTAAGTCAGTTTGA 321
QY      380 ThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGlu 399
Db      322 GACCGTAAAAAACAACAAAAATATTGATGGGCTCGACTTAACTGTTGGTTTCACGGGAAG 381
QY      400 LeuThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLys 419
Db      382 TTCACCTTAAAGTGAATTACAG- - - - - - - - - - -CAGCTTAAGCCCGT 417
QY      420 AspArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPhePro 439

```

Db 418 GAACGTAATTCCTGAGTTTCGACACGCAACACAGCTTATAATGACCTTTACCTGTCCTCA 477
Qy SerLeuLysMetValLeuGluSerLeuProGluasp-----Val 452
Db ACTCTAGAAACAATCATTCAGCTTCGGAAGCTAACTATAAAAGACGGGGAATAATA 537
Qy GlyPheAsnIleGluLeuLysTrpIleCysGlnArgAspGlyMetTrpAspGlyAsn 472
Db GGTATTATATTAGACGAAACAT-----561
Qy LeuSerThrTyrPhe---AspMetAsnLeuPheLeu---AspIleIleLeuLysThrVal 490
Db ---CCGACTATTTTAAATAATCAAAATCTGGCAATGGAAGACTACTCTCTTAAAAACCTTA 618
Qy LeuGluAsnSerGlyLysArgIleVal-----PheSerSerPheAsp-----505
Db GCCAAATATAATATACACGTGATATTGCACCTGTCTATTACAGCTCTTTTGAAGTTCAG 678
Qy 506 -----AlaAspIle 508
Db 679 AATCTAAAGATTATAAAGGGAGCTTGACCTTCATAAGACGCTTAAACACGCAAAAT 738
Qy CysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeuThrGlnGlyLys 528
Db 739 ATTACGCTATAGACTCAAAACATCTCGA---CCAGCAGACTTCGTAGAGCTGTGTGAC 795
Qy SerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThrProIleAlaMetSer 548
Db 796 ACTAAACCTTATGCTGATT-----GCCACAGCACCAAGGGTTAAAGAT 840
Qy PheAlaGlnPheGluAsnLeuLeuGly-----IleAsnValHisThr 562
Db 841 GTTGCCAAATATGCAAAATGGTGTAGGACCAAGTAAAGTTTACATACGTACCTTTAATAAT 900
Qy GluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysGlyLeuValIle 582
Db 901 GATGGCTCTTATAAACTAGTACGTTTATTCGTATGACACATACGCGCTGGTTTAAAGTA 960
Qy 583 PheCysTrpGlyAspAspThrAsnAspProGluAsnArg-----595
Db 961 CATCCTTAT-----ACTTCCGACCAGAAACAACTTCTACCAGCGCGTTAAG 1011
Qy 596 -----ArgLysLeuLysGluLeu-----601
Db 1012 TGCAGCCAGATAAACTGCTGAACTGTGTCCTCAACTGGTGGCTTAAAGAGTTTGAAGC 1071
Qy 602 -----GlyValAsnGlyLeuIleTyrAsp 609
Db 1072 TATTTCAAGCGAGGTGTTGATGCGCTTTTACCGAT 1107

RESULT 5

US-09-328-352-1148
; Sequence 1148, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1148
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1148

Alignment Scores:
Pred. No.: 2,22e-08 Length: 720
Score: 153.00 Matches: 67
Percent Similarity: 37.11% Conservative: 41
Best Local Similarity: 23.02% Mismatches: 115

Query Match: 4,33% Indels: 68
DB: 4 Gaps: 9
US-10-047-855-3 (1-672) x US-09-328-352-1148 (1-720)
Qy 321 ValGlyHisArgGlyValAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
Db 16 ATTGTGTCATCGTGGTGCACCGCA-----GAAGCTCTGAA 51
Qy 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaPheValGluPheAsp 360
Db 52 AATACATAGGTGGTTCAGCACATATAAAATTTAGGTATTTCGTGCTCGAATTTGAT 111
Qy 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
Db 112 ATTCTGTCAGCTCAAGATCAAGAGTTGGTTCGTTATTTCATGATGATAAATTTTTCGCGCAG 171
Qy 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
Db 172 GCA-----GGTGTCCGACCAAAATTTAGAGCACTCAACATTTGGCTCAAGCATTA 219
Qy 401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
Db 220 TCATTTGATCATCGT-----234
Qy 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
Db 235 -----CAAAATTTGGCCTAACTGGCCAACTCTGACCAACCCCTACT 276
Qy 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluLeuLysTrp 460
Db 277 CTTCAGAGTGTGTAAATTTACTTGATAACTTTTGATCATATTGAAAGTTGAAGTTAAAGCT 336
Qy 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
Db 337 GTA-----AGAGATATGGCA 351
Qy 481 LeuPheLeuAspIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal 500
Db 352 TTAGCAGAAAAAGTTAGTTCAAAAGCTTGAGACTGAGTTACAAGGTTTCGAGAGAAAGTCGTG 411
Qy 501 ---PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTrp 519
Db 412 ACCATCAACAAGTTTTCATCTTCAAAATTTCAACCGCTTTACGTCATTAATTAACCTCAATTT 471
Qy 520 ProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArg 539
Db 472 -----AAACGCGGATTTGTTAGTTGAACCT-----495
Qy 540 SerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGly---Ile 558
Db 496 -----CCGTTTGGAGCAACGGCGATTGAACCTGCTCATCAATATGCGTCTGT 543
Qy 559 AsnValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLys 578
Db 544 CATATTGGTGGCATGATCAACTCAACAGATGAATGATTCATTAATTTGCTCATCAAGCT 603
Qy 579 GlyLeuValIlePheCysTrpGlyAspThrAsnAspProGluAsnArgArgLysLeu 598
Db 604 GGTTTGAACATTAGTGTCTGG-----ACAGTGAACGATGTAGAAAGAGCAAGAGACTA 657
Qy 599 LysGluLeuGlyValAsnGlyLeuIleTyrAsp 609
Db 658 CAGGATTATGATATTTCAGGGCTTATTTACAGAT 690

RESULT 6

US-09-902-540-1012
; Sequence 1012, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

```

; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1012
; LENGTH: 14570
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1012

Alignment Scores:
Pred. No.: 3,83e-05 Length: 14570
Score: 145.00 Matches: 112
Percent Similarity: 33.64% Conservative: 70
Best Local Similarity: 20.70% Mismatches: 173
Query Match: 4.10% Indels: 187
DB: 4 Gaps: 27

US-10-047-855-3 (1-672) x US-09-902-540-1012 (1-14570)
Qy 136 ArgLeuHisTyrSerGluLysProValSerLeu-----ThrLysLysLysLeu-Ly 153
Db 4941 CGCTCGCATGGACAGCGAGGCGACCCCGCTCATCGCTGGATCGAAGAGAGGCGACCG 5000
Qy 153 slySerArgPheArgValLysLeuThrLeuGlyLeuGluGluAspAspAspA 173
Db 5001 CCAGGCGCGTGACG-----CAGGACGTGACCGCG 5033
Qy 173 gValSerProThrValLeuHisLysMetSerAsnSerLeuGluLeuSerLeu 193
Db 5034 ACCCTGGCTGCGTTCGGCAC-----TGCGCCAGTCGCGATTC 5072
Qy 193 pAnGluPheLysCysArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspA 213
Db 5073 CGGAGTCCAAACGGCGCGCCACGGTATCGGTTCAGTTGGATACGGACG-GCAACCCCTGG 5131
Qy 213 gTTPThrGluTyrSerLeuGlnThrMetGluProAspAsnLeuLeuLeuPheAspPh 233
Db 5132 GTGG----- 5135
Qy 233 ePheGluGluAspLeuSerGluHisValValGlnGlyAspAlaLeuProGlyHisVal 253
Db 5136 -----CATGGATGCCAGCCCTGGCGGTCCCTCCAC----- 5168
Qy 253 yThrAlaCysLeuLeuSerSerThrLeaGluSerGlyLysSerAlaGlyLeuTh 273
Db 5169 -----GAATCTTCGTCTACCGCTTCAACCGCTAG 5197
Qy 273 rLeuProIleMetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyr 293
Db 5198 CCAGCCCCACCGCGCC-----AGTCGTACACTGGCGGC-----AT 5233
Qy 293 eIleIleLysProLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyr 313
Db 5234 GGCATCTCTCCCTCC-----TTCTCCGCGCGCT 5266
Qy 313 pLysProArgIleProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAl 333
Db 5267 GCGGCCC-----ACGTGACATCGCCACCGCTGGGGCGCG----- 5303
Qy 333 aGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaA 352
Db 5304 -----GCCGTGGCGCGGAGAACGCTGGCGGCTTCGGCAGCGCTGGAGCGCTA 5356
Qy 352 sGlyAlaAlaPheValPheAspValHisLeuSerLysAspPheValProValVal 372
Db 5357 TCGCAGCGACATGTGGAGCTCGACCTCCACTCACCGGCGGGAGCTCGTGTGC 5416
Qy 372 rHisAspLeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspProValGlu 392

```

RESULT 7

```

US-09-543-681A-2048
; Sequence 2048, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A

```

```

5417 CCATGAC-----GCCACCTGGAGCGCTGCACGCGACGCGACGCGC----- 5456
Qy 392 uPheGluIleProValLysGluLeuThrPheAspGlnLeuGlnLeuLysLeu---Th 411
Db 5457 -----CCTCTGGCGCGCTCAGCTGCGCGAATCTCCAGCGCGCTGGACGCGGCTT 5506
Qy 411 rHisValThrAlaLeuLysSerLysArgLysGluSerValValGlnGluLys 431
Db 5507 CCACCTTCACGCCCGATGAAGCGC----- 5531
Qy 431 rPheSerGluAsnGlnProPhe-----ProSerLeuLysMetVa 444
Db 5532 -----ACCCTCCCATTCGGGGCCAGGCGCGTGCATCCCGAGCTTCGCGAATC 5581
Qy 444 lLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGln 464
Db 5582 GCTGCGCGCTTCCCC---AACCTGGCGCTCAACGTGGAGCTCAG----- 5624
Qy 464 nArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPhe 484
Db 5625 -----CCGGATGTCCCGGCATCGAGGA 5647
Qy 484 pIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSer 504
Db 5648 CACCTTCGCCCGAGTGTTCAGAGGAGAGCGCCCTGGAGCGGCTGTGATGGGCGAC-- 5705
Qy 504 eAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeu 524
Db 5706 -----GAATGGACACCGTGGCGGCGGTGGCGGAGCGTCTCCCT----- 5747
Qy 524 uThrGlnGlyLysSerGluIleTyrProGlu-----LeuMetAspLe 538
Db 5748 -----TCCGCTGCCACTTCTACCCCGCGATGCATCGCGCGCTTCTCATCGCCT 5800
Qy 538 uArgSerArgThrPro-----IleAlaMetSe 548
Db 5801 GCGTGGCGGAGACACGCGCGCGGAGGACCGCGCTACACCGTGTCTGACATGCCGCTCTA 5860
Qy 548 rPheAlaGlnPheGluAsnLeuGlyIleAsnValHisThrGluAspLeuLeuArg 568
Db 5861 CTTCCGTGAGATCCGG-----CTGGTGA 5884
Qy 568 nProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpGlyAsp 588
Db 5895 CTCGACTTCTCTCCAGCAGTCCGCGCGCGGCAAGTGGGTCAACGCTCTGACGCGTGA 5944
Qy 588 pThrAsnAspProGluAsnArgArgLysLysLysGluLeuGlyValaGlyLeuIleTy 608
Db 5945 T-----GACCGCGCGGAGATGCACCGCTCTCGCGGAGGCGTGGCGGCATCATGAC 5998
Qy 608 rAspArgIleTyrAspTrpMetProGluGlnProAsnIlePhe---GlnValGluGln 627
Db 5999 CGACCGG-----CCGACGCTCTGAGGCGAGATCATGACGCGC 6034
Qy 627 uGluArgLeuLysGlnGluLeuProGluLysSerCysLeuCysProThrValSerAr 647
Db 6035 CCCCTCGAAGCGGGATAAGCCCCGAT-----TCATGCTCGCACCAACCGCTCGCAG 6088
Qy 647 g 647
Db 6089 C 6089

```



```

; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 71
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 148..1140
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 148..240
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 10
; OTHER INFORMATION: seq LVLLVTRSPVNA/CL
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1590..1595
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 1614..1629
; US-09-247-155-71

Alignment Scores:
Pred. No.: 3,7e-05 Length: 1629
Score: 131.50 Matches: 82
Percent Similarity: 36.95% Conservative: 61
Best Local Similarity: 21.19% Mismatches: 106
Query Match: 3.72% Indels: 138
DB: 3 Gaps: 20

US-10-047-855-3 (1-672) x US-09-247-155-71 (1-1629)

QY 245 GlyAspAlaLeuProGly-HisValGlyThrAlaCysLeuLeuSerThrIleAlaG1 264
DB 216 GGTGACGGAGCCCGGTCAAT- - - - - GCTGTCTCTCACCAGCGCTC- - - - - 261
QY 264 userGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsnSerArgLysTh 284
DB 262 - - - - - TTGTTCTTACTGCGGTCTTCAGC- - - - - 285
QY 284 rIleGlyLysValArgValLeuThrIleIleLysProLeuProGlyTyrSerCysAs 304
DB 286 - - - - - TTGAGCCGGTGCC- - - - - TCTTGC- - - - - 306
QY 304 pMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp- - - ValGlyHi 323
DB 307 - - - - - AGGGCCCTGCAGGTGCTCAAGCCCGGGACCGCATTTCTGCCATCGCCCA 356
QY 323 sArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAsnThrI1 343
DB 357 CGTGGCGGACCCACGAC- - - - - GCGCCCGGAGAACACGCT 392
QY 343 eAlaSerLeuArgAsnAlaLaSerHisGlyAlaAlaPheValGluPheAspValHisLe 363
DB 393 GCGGCCCATTCGGCAGCGAGCTAAGATGAGAGCAACAGCGCTGGAGTTGGACATTGAGTT 452
QY 363 uSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMetLysLy 383
DB 453 TACTTCTGACGGGATTCCTCTTAAATGACGAT- - - - - AACACAGTAGATAG 500
QY 383 sLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAs 403
DB 501 GACGACTGATGGAGCTGGG- - - - - CGATTGTGTGATTGACATTGA 542
QY 403 pGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysG1 423
DB 543 ACAATATTAGAGCTGAATCTCTGACGAAACCCAGACATCCTGAGCAAGCTCAGG- - - - - 585

363 userLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMetLysLy 383
405 TACTTCTGACGGGATTCCTCTTAAATGACGAT- - - - - AACACAGTAGATAG 452
QY 383 sLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAs 403
DB 453 GACGACTGATGGAGCTGGG- - - - - CGATTGTGTGATTGACATTGA 494
QY 403 pGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysG1 423
DB 495 ACAATATTAGAGCTGAATCTCTGACGAAACCCAGACCTCAGG- - - - - 537
QY 423 userValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMe 443
DB 538 - - - - - NATGATTCCTCGAT- - - - - GAAAAGATCCCTACCTTAAGG- - - 573
QY 443 tValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysG1 463
DB 574 - - - - - GAAGCTGTTGCAGAGTGCCTAAACCATACCTCAATC- - - - - 612
QY 463 nGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMet- - - - - 479
DB 613 - - - - - TTCTTTGATGTCAAAGCCCATGC 635
QY 480 -AsnLeuPheLeuAspIleIleLeuLysThrValLeuGlu- - - - - As 493
DB 636 ACACAGGCTACTGAGGCTCTAAAGAAATGATATGGAATTCCTCAACTGTATAATAA 695
QY 493 nSerGlyLysArgArgIleValPheSerSerPhe- - - - - 504
DB 696 TAGT- - - - - GTGCTCTGTTCTTTCTTTCGCGAGAGTATTATCTACAAGATGAG 740
QY 505 -AspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLe 522
DB 741 ACAACAGATCGGGATGTAATACAGCATTAACCTACAGCA- - - - - CCTTGGAGCCTAAG 794
QY 522 uPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgTh 542
DB 795 CCATACAGGAGATGGGAACCA- - - - - CGCTATGATAC 827
QY 542 rThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisTh 562
DB 828 TTTCTCGAAACATTTTATATTTGTTATGATGAGCATTTTGTCTGATGGAGCATATAA 887
QY 562 r- - - - - GluAspLeuLeuArgAs 568
DB 888 TATCTTTGTGTACTGTGTGGAATTTACGTTTCTCATGCAAAAGATTTTGTGA- - - TC 944
QY 568 nProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpGlyAspAs 588
DB 945 CCGGCCCTACTGAAGAGTGGTCAGCTAAAGGATCCAGGTGTTGTTGGTGGACTGTAA 1004
QY 588 pThrAsnAspProGluAsn 594
DB 1005 TACCTTTGATGAAAGAGT 1023

RESULT 9
US-09-247-155-71
; Sequence 71, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; TITLE OF INVENTION: Bouqueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13

```

QY 423 uSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMe 443
Db 586 -----AATGATTTC-----CCTGATGAAAGATCCCTAC 614
QY 443 tValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGI 463
Db 615 CCTAATGGAAGCTGTTCAGAGTGCTTAACCACTAATCCTCACAATC----- 660
QY 463 nGlnA-gAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMet----- 479
Db 661 -----TTCTTTGATGTCAAAGGCCATGC 683
QY 480 -AsnLeuPheLeuAspIleIleLeuLysThrValLeuGlu-----As 493
Db 684 ACAACAGGCTACTGAGGCTCTAAGAAATGTATATGGAATTTCTCAACTGTATAATAA 743
QY 493 nSerGlyLysArgIleValPheSerSerPhe----- 504
Db 744 TAGT-----GTGCTCTGTCTTTCTTTCGCCAAGATTATCTACAAGATGAG 788
QY 505 -----AspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLe 522
Db 789 ACAACAGATCGGATGTAATAACAGCATTAACCTACAGA-----CCTTGGAGCCTAAG 842
QY 522 uPheLeuThrGlnGlySerGluIleTyrProGluLeuMetAspLeuArgSerArgTh 542
Db 843 CCATACAGGAGATGGGAAACCA-----CGCTATGATAC 875
QY 542 rThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisTh 562
Db 876 TTTCTGGAACATTTTATATTTCTTATGATGAGCATTTTCTCGATGGAGCATGATAA 935
QY 562 r-----GluAspLeuLeuArgAs 568
Db 935 TATCTTGTGTACTGTGTGAATTTTCAGCTTTCCTCATGCAAAAGATTTTGTA---TC 992
QY 568 nProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpGlyAspAs 588
Db 993 CCGGGCTACTTGAAGAAGTGGTCAGCTAAAGAAATCCAGGTGTGTGGTGGTACTGTAA 1052
QY 588 pThrAsnAspProGluAsn 594
Db 1053 TACCTTTGATGAAAGAGT 1071

RESULT 10

US-09-105-697-9
; Sequence 9, Application US/09105697
; Patent No. 6228628
; GENERAL INFORMATION:
; APPLICANT: Gelfand Ph.D., David H.
; APPLICANT: Reichert, Fred L.
; TITLE OF INVENTION: MUTANT CHIMERIC DNA POLYMERASE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Roche Molecular Systems
; STREET: 1080 U.S. Highway 202
; CITY: Branchburg
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 08876
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,697
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Petry Ph.D., Douglas A.
; REGISTRATION NUMBER: 35321
; REFERENCE/DOCKET NUMBER: 1043

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510)814-2974
; TELEFAX: (510)814-2977
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2682 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-105-697-9

Alignment Scores:

Prod. No.: 0.000203 Length: 2682
Score: 128.50 Matches: 140
Percent Similarity: 31.91% Conservative: 69
Best Local Similarity: 21.37% Mismatches: 195
Query Match: 3.63% Indels: 251
DB: 3 Gaps: 34

US-10-047-855-3 (1-672) x US-09-105-697-9 (1-2682)

QY 150 LysLysLeuLysLysSerArgPheArgValLysLeuThrLeuGluGlyLeuGluAsp 169
Db 379 AAGAAGCGGAAAGAGGGCTACGAGGTCCGCATC-----CTCACCGCCGAC 426
QY 170 AspAsp-----AspArgValSerProThrValLeuHisLysMetSer 183
Db 427 AAGACCTTTTACCAGCTCTTCCGACCCGATCCAC-----GTCTCCACCCGAGGGG 480
QY 184 AsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHisSerGlnPro 203
Db 481 TACCTCATCACCCGCGCTGGCTTTGGGAAAG----- 513
QY 204 GluCysGlyTyrGlyLeuGlnProAspArgTTrpGluTyrSerIleGlnThr---Met 222
Db 514 -----TACGGCTGAGGCCCGACCTACGAGTGGCGGCTACCGGCCCTGACCGGGGAC 564
QY 223 GluProAspAsnLeuGluIlePheAspPheGlu----- 235
Db 565 GAGTCCGACACATCCCCGGGTCTACTGGGATCGGTGAGAAAGACTGCTGTCAGCTTCTA 624
QY 236 -----GluAspLeuSerGluHisValGlnGlyAspAlaLeuPro 249
Db 625 GAGAAGTACAAAGACCTCGAAGACATACCTGAATCATGTTCCGGAA-----CTTCT 675
QY 250 GlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAlaGluSerGlyLysSerAla 269
Db 676 CAAAAGGTGAGAAAGCCCTGCTTCGAGACAGAGAAACGCCATTCTCAGCAAAAAGCTG 735
QY 270 GlyIleLeu-----ValArgValAspTyrIleIleIleLysProLeu--- 298
Db 736 GCGATTCTGGAACAAACCTTCCCATTTGAATAAATCTGGGAAAGAACTTCCTACACAGGC 795
QY 273 -----ThrLeuProIleMetSerArgAsnSerArgLysThrIleGlyLys 287
Db 796 TAGCAGACAGAGAAACTCTTACCACCTTTTGAAGAACTCGAATTCGCATCCATCATGAAG 855
QY 288 -----ValArgValAspTyrIleIleIleLysProLeu--- 298
Db 856 GAACTTCAACTGTACGAGAGTCCGAACCCGTTGGATACAGAAATAGTGAAGACCTTAGT 915
QY 299 -----ProGlyTyrSerCysAspMetLys 306
Db 916 GAATTTGAAAAAATCATAGAGAACTGAGAGAAATCCCTTCGTTCCCATATAGATCTTAG 975
QY 307 SerSer-----PheSerLysTyrTrpLysPro 315
Db 976 ACGTCTCTCCCTCGATCCTTTCCAGCTGCGACATTGTCGGTATCTCTGTGCTTTCAACCA 1035
QY 316 Arg-----IleProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThr 331
Db 1036 AAGGAAGCGTACTACATACCACTCCAT-----CATAGAAACGCCCAAGACCTGGACGAA 1089


```

QY 332 ThrAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSer 351
Db 1090 AAGAGGGTTCGAAAAAGCTCAAGAA-----ATTCTGGAGGAC 1128
QY 352 HisGlyAlaAlaPheVal-----GluPheAspValHisLeuSerLys 365
Db 1129 CCGGAGCAAAAGATCGTTGGTCAGAAATTTGCAATTCACAAAGGTGTGATGGTGAAG 1188
QY 366 AspPheValProVal---ValTyrHisAspLeuThrCysCysLeuThrMetLysLysLys 384
Db 1189 GGTGTGAACCTGTTCTCTCTTACTTCAC-----ACGATGATAGCGCT 1233
QY 385 PheAspAlaAspProValGlnLeuPheGluLeuProValLysGluLeuThrPheAspGln 404
Db 1234 TACCTCTTGGAGCGAAGCA-----AAGAAGTTTCAATCTGGAGCAT 1275
QY 405 LeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSer 424
Db 1276 CTC-----GCATTGAAATTTCTTGGATACAAATGACA 1308
QY 425 ValValGlnGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMetVal 444
Db 1309 TCTTACCAGAGCTATGTCCTCTCT-----TTTCGCTGTGGTTTCAGT 1356
QY 445 LeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGlnGln 464
Db 1357 TTTGCCGATGTCCTGTAGAAAAAGCAGCGAAC-----TACTCTGTGAAGAT 1404
QY 465 ArgAsp-----GlyMetTrpAspLysLeu 473
Db 1405 GCAGACATCACCTACAGACTTTTACAGACCTGAGCTTAAACTCCAGAGCAGCATGTG 1464
QY 474 SerThrTyrPhe-----AspMet 479
Db 1465 GAAAAAGTGTCTTACAGATAGAAATGCCCTTGTGAACGTGTCACGGATGGAACGTG 1524
QY 480 Asn---LeuPheLeuAspIle---IleLeuLysThrValLeuGluAsnSerGlyLysArg 497
Db 1525 AACGGTGTGTATGTGGACACAGAGTTCTCTGAAGAACTCTCAGAGAGATGACGAAAAAA 1584
QY 498 -----ArgIleValPheSerSerPheAspAlaAspIle 508
Db 1585 CTCGAGAACTGGCAGAGGAAATATACAGATAGCTGGAGAGCGGTTCACACTAACTCA 1644
QY 509 CysThrMetVal-----ArgGlnLysGln 516
Db 1645 CCGAAGCAGGTTTCAAGGATCCTTTTGAAGAACTCGGCATAAACCACTGGTAAACG 1704
QY 517 AsnLys-----TyrProIleLeuPheLeuThrGlnGlyLysSer 529
Db 1705 ACGAAACGGGAGACTATTCAACACGCATAGAGTCTCTCAGGAACTTGCCTGGAACAC 1764
QY 530 GluIleTyrProGluLeuMetAspLeuArgSer----- 540
Db 1765 GAAATCATCTCTGATTTTGAATACAGAAAGATTCATGCTTCTTCAATCAA 1884
QY 541 -----ArgThrThrProIleAlaMetSerPheAlaGln 551
Db 1825 GACGCTCTTCCCAAGATGTCACCCCAAGHCCGGAAGGATTCATGCTTCTTCAATCAA 1884
QY 552 PheGluAsnLeuLeuGlyIleAsnValHisThrGluAspLeuLeuArgAsn---ProSer 570
Db 1885 ACGGGAGTCCACTGGAGACTTAGCAGCAGCGATCCCAATCTTCAAGAACCTCCCGCAGC 1944
QY 571 TyrIleGlnGluAlaLys-----AlaLysGlyLeuValIlePheCysTyrTrpGlyAsp 588
Db 1945 AAAAGTGAAGGGGAAAGAAATCAGGAAACCGATAGTT----- 1983
QY 589 ThrAsnAspPro----- 592
Db 1984 CCTCAGGATCCAAACTGGTGGATGCTCAGTCCGACTACTCCCAATAGACTGAGGATC 2043
QY 593 -----GluAsnArgArgLysLeuLysGluLeuGly-----Val 603
Db 2044 CTCGCCCATCTCAGTGGTGCATCAGAATCTTTTGGAGGCATTTGGAAGGGCATTCGACGTC 2103
QY 604 AsnGlyLeuLeuTyrAspArgIleTyrAspTrpMetProGluGln 618
Db 2104 CACACTCTAACAGCTTCCAGAAATATTCAACGTGAACCCGAGAA 2148

```

RESULT 11

US-09-252-991A-1900

; Sequence 1900, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 1900

; LENGTH: 1173

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-1900

Alignment Scores:

Pred. No.: 8,45e-05 Length: 1173

Score: 126.50 Matches: 87

Percent Similarity: 36.55% Conservative: 53

Best Local Similarity: 22.72% Mismatches: 146

Query Match: 3.58% Indels: 98

DB: 4 Gaps: 17

US-10-047-855-3 (1-672) x US-09-252-991A-1900 (1-1173)

QY 244 GlnGlyAsp-----AlaLeuProGlyHisValGlyThr-----AlaCysLeuLeu 258

Db 227 CAGGCGCATATACCCGGGTGCGCTCCCGCGCAGCGGAAAGCCCGCGGAGCGTCTGCAA 286

QY 259 SerSerThrIleAlaGluSerGlyLysSerAlaGly-----IleLeu 272

Db 287 CGCTTCGCGCTGGCGGACGA--GGAACATCCCGCGTCCGCGCAGCGCCATCGTGGCTG 345

QY 273 ThrLeuProIleMetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTyr 292

Db 346 GGATTTACCGCAGCTTCTCGGTATATTTGAAAGCGCGCGGTGCTCGACAGCGTTGCG--- 402

QY 293 IleIleIleLysProLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyr 312

Db 403 -----CCATCATCGCGCGCGGAAATCCCGCGCAACCTCGAACAA--- 441

QY 313 TrpLysProArgIleProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrThr 332

Db 442 ---GGAGCCACAGTGACTCTCATCTATGGACATCCGCGCGCC----- 480

QY 333 AlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHis 352

Db 481 -----AAGGGCGAAGCGCGGAAACACCTCGGCAGCTTCCAGCGCTGCTGAGCAT 534

QY 353 GlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheValProValValTyr 372

Db 535 GGGGTCAACCGTTGCGAAGTGGACCTGACCTGCTCCCGCGCAGCGGCGGAACTGATGTGATC 594

QY 373 HisAspLeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspProValGluLeu 392

Db 595 CAGACCCGACCC----- 606

QY 393 PheGluIleProValLysGluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHis 412

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1707
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1707

Alignment Scores:
Pred. No.: 8,19e-05 Length: 771
Score: 124.00 Matches: 63
Percent Similarity: 37.04% Conservative: 47
Best Local Similarity: 21.21% Mismatches: 113
Query Match: 3.51% Indels: 74
DB: 11 Gaps: 11

US-10-047-855-3 (1-672) x US-09-134-001C-1707 (1-771)

QY	322	GlyHisAArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAen	341
DB	43	GGACATAGAGGATTACCT-----AGTAAAGCTCCGGAAAT	78
QY	342	ThrIleAlaSerLeuArgAsnAlaAlaSer---HisGlyAlaAlaPheValGluPheAsp	360
DB	79	ACAATTGCATTTAAAGCTGCTCAGAAAGTAGAAGGTATAAAGCTGGTGGAGTTAGAT	138
QY	361	ValHisLeuSerLysAspPheValProValValTyrHisAsp-----LeuThrCysCys	378
DB	139	GTTGCAATTACAAAGATGAAACAACTGATATCATTCATGATGATATTTAGAACGGACT	198
QY	379	LeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLys	398
DB	199	ACMAATATGTCCGGGAA-----ATAACT	222
QY	399	GluLeuThrPheAspGlnLeuGlnLeuLysLysLeuThrHisValThrAlaLeuLysSer	418
DB	223	GAATTGAAATTATGATGAATTTAAAGATGCTTCTGCAGGATCTTGGTTGGTGAAATATTC	282
QY	419	LysAspArgLys-----GluSerValValGlnGluGluAsnSerPheSerGlu	434
DB	283	AAAGATGAACATTTGCCAATTTCCGATGATGATGATGATGATGATGATGATGATGATG	342
QY	435	AsnGlnProPheProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPhe	454
DB	343	AAAT-----TTA	348
QY	455	AsnIleGluIleLysTrpIleCysGlnArgAspGlyMetTrpAspGlyAsnLeuSer	474
DB	349	AATGTAGATTTAAAGGTTAT-----ACTGACCGAATGACTAGCA	390
QY	475	ThrTyrPheAspMetAsnLeuPheLeuAspIleIleLysThrValLeuGluAsnSer	494
DB	391	CTTTCTAAAGTATGGTTAAGCAAGTGGGAAGAACAAATTTACAACTTAATCAGAT---	447
QY	495	GlyLysArgArgIleValPheSerSerPheAspAlaAspIleCysThrMetValArgGln	514
DB	448	-----CAAGAGTGCTCATTTCAAGCTTTAATGTTGTGCTTGTAACTTCGAGAAGAA	501
QY	515	LysGlnAsnLysTyr-----ProIleLeuPheLeuThrGlnGlyLysSerGluIleTyr	532
DB	502	ATCATGCCACAATATAACAGACGAGGTATATTTCCATACAACTTCGTTTCGTGAAGACTGG	561
QY	533	ProGluLeuMetAspLeuArgSerArgThrThrProIleAlaMetSerPheAlaGlnPhe	552
DB	562	AGAACACTTTTAGATTACTGTAATGCTAAATA-----	594

QY	553	GluAsnLeuLeuGlyIleAsnValHisThrGluAspLeuLeuArgAsnProSerTyrIle	572
DB	595	-----GTAACACTGAAGATGCCAAACTTACTAAAGCAAAAGTA	633
QY	573	GlnGluAlaLysGlyLeuValIlePheCysTrpGlyAspThrHisAspPro	592
DB	634	AAATGGTAAAGAGAGCGGTTATGAATTGAACGTATGG-----ACTGTAACAAACCA	687
QY	593	GluAsnArgLysLeuLysGluLeuGlyValAsnGlyLeuIleTyrAsp	609
DB	688	GCACGTGCAACCAACTTCTAATTGGGGAGTTGATGGTATCTTTACAGAC	738

RESULT 14

US-09-902-540-1212/c
; Sequence 1212, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1212
; LENGTH: 26012
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1212

Alignment Scores:
Pred. No.: 0.0369 Length: 26012
Score: 124.00 Matches: 70
Percent Similarity: 36.91% Conservative: 40
Best Local Similarity: 23.49% Mismatches: 108
Query Match: 3.51% Indels: 80
DB: 13 Gaps: 13

US-10-047-855-3 (1-672) x US-09-902-540-1212 (1-26012)			
QY	321	ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu	340
Db	24817	CTTGCCACCGTGGCGCCAGC-----GCCGATGCCCGCCGAG	24782
QY	341	AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp	360
Db	24781	AACACCTGGAGCCTTCGCCGAGCGGTCCGCCAAGCGCGGAGCGGTGTGGAACCTGAC	24722
QY	361	ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr	380
Db	24721	GCATGCTCTGTGGCTCGCGCGAGGTCTGCTGTGCCACGAC-----	24680
QY	381	MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu	400
Db	24679	-----GAGCGCCTGGAG-----CGCCTCGCGCGGCTGCCGTGGAG-----	24644
QY	401	ThrPheAspGlnLeuGlnLeuLysLysLeuThrHisValThrAlaLeuLysSerLysAsp	420
Db	24643	-----GTTCCGCTCACCCCGTGTGGAGAGCTCCAGCGGCGGAC	24605
QY	421	ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer	440
Db	24604	-----GTCGGCTCGCGCTGGGCTTCGCCCGCCGCCCGA---ATCCCGCTG	24563
QY	441	LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp	460
Db	24562	CTGGAAGGTGCTGGAGGCGCTCCACCGCAGATTCCTGTCACATCGAGCTCAAG---	24506
QY	461	IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn	480

817 TAAGTTAAAGGAAGATTATGAAAAATCCACACCTTGAACAGAATACACAGAGGAAT 876
Db
248 uProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAlaGluSerGlyLysSe 268
Qy
877 AAATGACAAAGAAAGCAGGTATCACTACTATTGATCCAAATCACTGAGAAAGAAAATAA 936
Db
268 rAlaGlyIleLeuThrLeu 274
Qy
937 AATGAAAGATTAAACATTTCTGCTAGAGGAATCCAGAGATAAAGTTAATCAATTAGAGGA 996
Db
275 -----ProIleMetSerArgAsnSerArgLysThrIleGlyLysValArgValAspTy 292
Qy
997 AAAGACAAAATACAGAGTGAACCTTAAACAAATCAATTGAGAAA-----CAGCA 1047
Db
292 rIleIleIleLysProLeuProGlyTyTrSerCysAspMetLysSerSerPheSerLysTy 312
Qy
1048 TCATTTGACTAAAGAACTAGAGATATTAAAGTGTCAATTACAAAGAGTGTGACTACTCA 1107
Db
312 rTrpLysProArgIleProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrTh 332
Qy
1108 AAAGCTTTAGAGGAAGATTACAGATAGCA-----ACAAAACAAT 1149
Db
332 rAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeuArgAsnAla---Alase 351
Qy
1150 TTGTGAGCTAACTGAAGAAAGAAAGAACTCAAAATGGAAGAACTAATAAGCTAGAGCTGC 1209
Db
351 rHisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAspPheValProValVa 371
Qy
1210 TCATTCGTTTGGTTTACTGAATTTGAACT----- 1240
Db
371 lTyHisAspLeuThrCysCysLeuThrMetLysLysLysPheAspAlaAspProValGl 391
Qy
1241 -----ACTGTCTGCAGCTTGAAGAATATTATTGAGACAGACAGCAAGAAAGATTGGA 1290
Db
391 uLeuPheGluIleProValLysGluLeuThrPheAsp-----GlnLe 405
Qy
1291 AAAAAATGAAGATCAATTTGAAAATCTTACCATTGAGCTTCAAAAGAAATCAAGTGAGCT 1350
Db
405 u---GlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSe 424
Qy
1351 GGAAGAGATGACTAAGCTTACAAATAACAAGAGTAGAAGCTTGAAGAAATGCAAAAAGT 1410
Db
424 rValValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMetVa 444
Qy
1411 CTGGGAGAAAAGAAACACTTTTATATGAAATAAACAATTT----- 1453
Db
444 lLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGlnGl 464
Qy
1454 ---GAGAAGATTGCTGAA-----GAATTAAAGGAACAGACAGCAAGA 1491
Db
464 nArgAspGlyMetTrpAspGlyAsnLeuSerThrTyPheAspMetAsnLeuPheLeuAs 484
Qy
1492 ACTAATTTGCTTCTCAAGCCAGAGAAAGAGTACATGATTTGGAAAATACAGTTAAC 1551
Db
484 pIleIle-----LeuLysThrValle 491
Qy
1552 TGCCATTACCACAGTGAACAGTATTATTCAAAAGAGGTTAAAGATCTAAAAACTGAGCT 1611
Db
491 uGluAsnSerGlyLysArgArgIleValPheSerSerPheAspAlaAspIleCysThrMe 511
Qy
1612 TGAAGACGAGAGCTTAAAGAAATACTGAATTAACCTCACAC-----TGCAACAA 1659
Db
511 tValArgGlnLysGlnAsnLysTyTrProIleLeuPheLeuThrGlnGlyLysSerGluI 531
Qy
1660 GCCTT---TCACAGAAAACAAGAG-----CTCACAGGAAACAAGTGTATAT 1704
Db
531 eTyTrProGluLeuMet-----AspLeuArgSerArgThrThrProIleAlaMe 547
Qy
1705 GACCCCTAGACTCAAGAATCAGCAAGAGATATTATAATAACAAAAAGCAAGAGAAAG 1764
Db
547 tSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGluAspLeuLeuAr 567
Qy
1765 GATGTTGAAAACAAATAGAAAAATCTT-----CAAGAAACAGAAACCCCAATTAAG 1812
Db

567 gAsnPro---SerTyIle---GlnGluAlaLysAlaLysGlyLeuValIlePheCysTr 585
Qy
1813 AAATGAACTAGATAATGTGAGAGAGAGCTAAAAACAGAAAAGAGATGAAGTTAAATGTAA 1872
Db
585 pGlyAspAspThrAsnAsp-----ProGluAsnArgArgly 597
Qy
1873 ATTGGACAAGAGTGAAGAAAAATTTGTAACAATTTTAAGGAAACAAGTTGNAATAAACAACAA 1932
Db
597 s---LeuLysGluLeuGlyValAsnGlyLeuIleTyAspArgIleTyAspTrpMetPr 616
Qy
1933 GTATATTGAAGAACTTTCAGCAGAGAGAAATAGGCCTTTGAAAAAAAAGGTACACGAGAAAG 1992
Db
616 oGluGlnProAsnIlePheGlnValGlu-----GlnLeuGluArgLe 630
Qy
1993 CAAGCAACTCAATGTTTATGAGATAAAGGTCAATAAATTAGAGTTAGAACTAGAAAAGTGC 2052
Db
630 uLysGlnGluLeuProGluLeuLysSerCysLeuCysProThrValSerArgPheValPr 650
Qy
2053 CAAACAGAAATTTGGAGAAATCACAGAC-----ACCTATCAGAAAGAAATTTGA 2100
Db
650 oSerSerLeuCysGlyGluSerAspIleHisValAspAlaAsnGlyIleAspAsnValGl 670
Qy
2101 GGCACAAAAGATATATCAGAAAGAAATCTT-----TTGGAAGAGGTTGA 2142
Db
670 uAsnAla 672
Qy
2143 GAAGCA 2149
Db

Search completed: July 3, 2005, 11:02:02
Job time : 352 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 3, 2005, 06:44:29 ; Search time 950 Seconds
(without alignments)
4187.438 Million cell updates/sec

Title: US-10-047-855-3

Perfect score: 3537
Sequence: 1 MTPSQVAFEIRGTLPLGEV.....LCGESDIHVANGIDNVENA 672

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DBV=xlh
-O/csp2_1/USPTO spool/US10047855/runat 01072005 155528 2208/app query.fasta_1.839
-DB=N_Geneseq 16Dec04 -QFMT=fastp -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10047855 @CN 1.1 644 @runat 01072005 155528 2208 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 16Dec04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001as:.*
5: Geneseq2001bs:.*
6: Geneseq2002as:.*
7: Geneseq2002bs:.*
8: Geneseq2003as:.*
9: Geneseq2003bs:.*
10: Geneseq2003cs:.*
11: Geneseq2003ds:.*
12: Geneseq2004as:.*
13: Geneseq2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	3537	100.0 2019 13	ADR40144 Human Nar
2	3537	100.0 3206 4	AAD06010 Human neu
3	3537	100.0 3206 8	ABS56723 Human NAR
4	3537	100.0 3206 12	ADJ57933 Human NAR
5	3537	100.0 3499 5	ABV29614 Human pro

6	3537	100.0	3499	5	ADL62788	AdL62788 Human ova
7	3537	100.0	5443	12	ADN05171	AdN05171 Antipsori
8	3537	100.0	5443	13	ADS74320	AdS74320 PRO83903
9	3509	83.8	3498	5	ABV23741	Abv23741 Human pro
10	2965.5	89.2	3381	12	ADJ57938	AdJ57938 Rat NARC
11	2829.5	80.0	2477	5	ABA09667	Aba09667 Human bon
12	2552	72.2	2738	4	AAD06007	Aad06007 Human neu
13	2552	72.2	2738	12	ADJ57931	AdJ57931 Human NAR
14	2439.5	69.0	1978	5	ABA09583	Aba09583 Human bon
15	1399	39.6	2393	12	ADJ57947	AdJ57947 Rat NARC
16	1364	38.6	1803	4	AAH14905	Aah14905 Human CDN
17	1216	34.4	1929	4	AAH44173	Aah44173 Human gly
18	1102.5	31.2	768	4	AAH07671	Aah07671 Human CDN
19	1005	28.4	2792	4	ABL05249	AbL05249 Drosophil
20	922	26.1	5439	4	ABL05248	AbL05248 Drosophil
21	824	23.3	471	6	ABV88538	Abv88538 Human col
22	717	20.3	520	5	ADL42779	AdL42779 Human ova
23	667	18.9	2392	4	ABL19961	AbL19961 Drosophil
24	667	18.9	2457	4	ABL19771	AbL19771 Drosophil
25	667	18.9	7162	4	ABL19770	AbL19770 Drosophil
26	667	18.9	9667	4	ABL19960	AbL19960 Drosophil
27	666	18.8	2299	4	ABL15151	AbL15151 Drosophil
28	666	18.8	4299	4	ABL15150	AbL15150 Drosophil
29	661	18.7	2220	4	ABL28903	AbL28903 Drosophil
30	661	18.7	4220	4	ABL28902	AbL28902 Drosophil
31	628	17.8	396	5	ADI71235	AdI71235 Human ova
32	628	17.8	396	5	ADI36395	AdI36395 Human ova
33	570	16.1	425	5	ABV35865	Abv35865 Human pro
34	570	16.1	425	5	ABV44927	Abv44927 Human pro
35	564	15.9	327	5	ABV15063	Abv15063 Human pro
36	550	15.5	2100	4	ABL24297	AbL24297 Drosophil
37	550	15.5	4100	4	ABL24296	AbL24296 Drosophil
38	482.5	13.6	3672	13	ADT47737	Adt47737 Bacterial
39	473	13.4	14598	4	AAK80690	Aak80690 Human imm
40	473	13.4	14598	4	AAK79627	Aak79627 Human imm
41	472.5	13.4	359	5	ABV05894	Abv05894 Human pro
42	442	12.5	4925	8	ADA22696	Ada22696 A. goesyp
43	440	12.4	474	4	AAK61234	Aak61234 Human imm
44	432	12.2	3799	13	ADS48391	AdS48391 Bacterial
45	430.5	12.2	2226	13	ADS60798	AdS60798 Bacterial

ALIGNMENTS

RESULT 1
ADR40144
ID ADR40144 standard; cDNA; 2019 BP.
XX
AC ADR40144;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human Narcl16b (64549) cDNA.
XX
KW haematological; cytostatic; erythroid; anaemia; erythrocytosis;
KW bone marrow; leukaemia; platelet; thrombocytopenia; thrombosis; B-cell;
KW T-cells; neutropenia; gene therapy; human; ss; gene; Narcl16b.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2019
FT /tag= a
FT /product= "Human Narcl16b (64549) protein"
XX
FN WO2004072242-A2.
XX
PD 26-AUG-2004.
XX
PF 05-FEB-2004; 2004WO-US003417.
XX
PR 05-FEB-2003; 2003US-0445241P.
PR 18-FEB-2003; 2003US-0448389P.

PR 20-MAR-2003; 2003US-0456320P.
PR 03-APR-2003; 2003US-0460279P.
PR 28-APR-2003; 2003US-0465924P.
PR 13-MAY-2003; 2003US-0470052P.
PR 26-AUG-2003; 2003US-0498106P.
PR 04-SEP-2003; 2003US-0500179P.
PR 15-SEP-2003; 2003US-0502909P.
PR 10-OCT-2003; 2003US-0510351P.
PR 17-OCT-2003; 2003US-0512380P.
XX (MILL-) MILLENNIUM PHARM INC.
PA
PI Kelly LM, Carroll JM, Farlow D, Healy A;
XX
XX WPI; 2004-625850/60.
DR P-PSDB; ADR40145.
XX
XX Identifying a compound capable of treating a hematological disorder
PT comprises combining a compound to be tested with a polypeptide related
PT with the disorder under conditions suitable for binding of the test
PT compound to the polypeptide.
XX
XX Disclosure; SEQ ID NO 25; 321pp; English.
XX
XX The invention relates to a novel method for identifying a compound
CC capable of treating a hematological disorder which comprises combining a
CC compound to be tested with a specific polypeptide under conditions
CC suitable for binding of the test compound to the polypeptide. The method
CC of the invention has hematological and cytostatic applications and may
CC be useful for identifying compounds for treating a hematological
CC disorder associated with erythroid cells e.g. anaemia and erythrocytosis,
CC bone marrow e.g. leukaemia, platelets e.g. thrombocytopenia and
CC thrombosis or B-cells and T-cells e.g. neutropenia. The compounds
CC identified may be utilised during gene therapy procedures. The current
CC sequence is that of a human hematological disorder-related cDNA of the
CC invention.
XX
XX Sequence 2019 BP; 633 A; 363 C; 449 G; 574 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2019
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-047-855-3 (1-672) x ADR40144 (1-2019)

QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
DB 1 ATGACACCTTCTCAGGTTGCCTTTGAAATAAGAGGAACCTTTTACCAGGAGAAGTTTTT 60
QY 21 AlaileCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
DB 61 GCGATATGTGGAAGCTGTGATGCTTTGGGAAACTGGAATCCTCAAAATCGCTGGCTCTT 120
QY 41 LeuProGlnAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
DB 121 CTTCCAGAGAAATGACACAGTGGAAGCATGCTATGGAAGCAACCACTTTGTACTCAGTAGA 180
QY 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
DB 181 GGAGTATCAGTTCAGTATCGTACTTCAAGGGGTACTTTTTAGAACCAAGACTATCGGT 240
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
DB 241 GGTCCATGTCAAGTGATAGTTTCAAGTGGGAGACTCATCTACAACCAACGATCAATAACC 300
QY 101 ProLeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
DB 301 CCTTTAGAAACCGAAATTTATTTATTTGACATGGACATTTGGATCCCAATGTTGTGAA 360

QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
DB ACTCTGGATTCTGGATGGCTGACATGTGCAGACTGAAATAAGATTACGTTTGCAATTATCT 420
QY 141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
DB GAAAAACCTCTCTGTGTCAATAACCAAGAAAAAATAAAAAAATCTAGATTTTAGGGTGAAG 480
QY 161 LeuThrLeuGluGluLysLeuGluAspAspAspArgValSerProThrValLeuHis 180
DB CTGACACTAGAAGGCCCTGGAGGAAGATGACATGATAGGGTATCTCCCACTGTACTCCAC 540
QY 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPhelLysCysArgHis 200
DB AANAATGTCANATAGCTTGGAGATATCTCTTAATAAGCGACAATCAGTTCAAGTGCAGGCAT 600
QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
DB TCACAGCCGAGTGTGCTTATGGCTTGCAGCCCTGATCGTTGGACAGATACAGATACAG 660
QY 221 ThrMetGluProAspAsnLeuGluLeuIlePheAspPhePheGluGluAspLeuSerGlu 240
DB ACCATGCTCAGAGTGGAAAGAGTGTGGAAATCTTACTCTTCCCACATGAGCAGAGAAAT 840
QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300
DB TCCCGGAAACAATAGGCAAGTGAGAGTTGACATATATATATTAATTAAGCCATTAACAGGA 900
QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
DB TACAGTTGTGNCATGAATCTTCAATTTTCCAAGTATTGGAAGCCAGCAATACCAATGGAT 960
QY 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
DB GTTGGCCATCAGGTCGAGGAAACTCTACAACAACATGCCAGCTGGCTTAAGTTCAAGAA 1020
QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
DB AATACTATTGCTCTTTAAGAAATGCTGCTAGTCTATGTCAGGCTTTGTAGATTTGAC 1080
QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysLeuThr 380
DB GTACACCTTTCAAGGACTTTTGTGCCGTGGTATATCATGATCTTACCTGTGTTGTGACT 1140
QY 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
DB ATGAAAAAGAAATTTGATGCTGATCCAGTCTGAATTTATTTGAAATTTCCAGTAAAAAGAAITA 1200
QY 401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
DB ACATTTGACCAACTCCAGTTGTTAAGCTCACTCATGTCAGTGCACCTGAATCTTAAGGAT 1260
QY 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
DB CGGAAAGAAATCTGTGTTGTCAGGAGGAAATCTCTTTTCAGAAAAATCAGCCATCTCTTCT 1320
QY 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
DB CTTAAGATGCTTTAGAGTCTTTTGCAGAGATGTAGGGTTTAAACATTGAAATAAAATGG 1380
QY 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
DB ATCTCCAGCAAGGATGGAATGTGGATGTGTAATCTTATCAACATATTTTGACATGAT 1440
QY 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal 500

Db 1441 CTTGTTTGGATATAATTTAAACAACTGTTTGGGAGGAGGAAATAGTG 1500
Qy 501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro 520
Db 1501 TTTTCTTCATTGATGCAGATATTGACANATGGTTGGCAAAAGCAGCAAAATATCG 1560
Qy 521 IleLeuPheLeuThrGlnGlySerGluLeuTyrProGluLeuMetAspLeuArgSer 540
Db 1561 ATACTATTTTAACTCAAGGAAAAATCTGAGATTTATCTGAACTCATGACCTCAGATCT 1620
Qy 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal 560
Db 1621 CGACAAACCCCAATGCAATGAGCTTTGCACAGTTTGAATAATCTACTGGGGATAAATGTA 1680
Qy 561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
Db 1681 CATACTGAGAGCTTGCTCAGAAACCATCTATATTCAGAGGCAAAAGCTAAGGACTA 1740
Qy 581 ValIlePheCysTyrPglyAspAspThrAsnAspProGluAsnArgLysLeuLysGlu 600
Db 1741 GTCATATCTCTGGGCTGATGATACCAATGATCTCTGAAACAGAAAGGAAATTTGAAGGAA 1800
Qy 601 LeuGlyValAsnGlyLeuLeuTyrAspArgIleTyrAspTyrMetProGluGlnProAsn 620
Db 1801 CTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATTGGATGCTCGAACCAACAAAT 1860
Qy 621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
Db 1861 ATATTCAGAGTGGAGCAATTTGAAACGCTTGAGCAGGAATTTGCCAGAGCTTAAGAGCTGT 1920
Qy 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 1921 TTGTGTCCCACTGTTAGCGCTTTGTGTCTCATCTTTGTGTGGGAGTCTGATATCCAT 1980
Qy 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 1981 GTGGATGCCACGGCAATTCATAACGTGGAGATGCT 2016

RESULT 2
AAD06010
ID AAD06010 standard; DNA; 3206 BP.
AC AAD06010;
XX 31-JUL-2001 (first entry)
XX Human neuronal apoptosis regulated candidate (NARC) 16B DNA.
XX Human; neuronal apoptosis regulated candidate 16B; NARC 16B; cytostatic;
KW chromosome mapping; gene therapy; antisense therapy; lung disorder;
KW central nervous system disorder; apoptosis; spleen disorder; angina;
KW tuberculous; Goodpasture's syndrome; liver disorder; jaundice;
KW infectious disorder; brain disorder; cerebral oedema; gonorrhoea;
KW heart disorder; kidney disorder; glomerulonephritis; testes; virucide;
KW epididymis disorder; skeletal muscle disorder; pancreatic disorder;
KW diabetes; cytoprotectant; immunostimulant; tumour; antiinflammatory;
KW antimicrobial; neuroprotective; gynaecological; ds.
XX Homo sapiens.
OS
XX
XX WO200131007-A2.
XX
XX 03-MAY-2001.
XX
XX 20-OCT-2000; 2000WO-US029132.
XX
XX 22-OCT-1999; 99US-0161188P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Chiang LW;
PI
XX

DR WPI; 2001-308641/32.
XX Rat brain polypeptides, nucleic acids and antibodies, useful for
PT diagnosis and treatment of central nervous system disorders and disorders
PT associated with aberrant apoptosis.
XX Claim 1; Page 148; 161pp; English.
XX The invention relates to human homologues of neuronal apoptosis regulated
CC candidate (NARC) nucleic acid molecules and proteins derived from rat
CC brain and programmed cell death libraries. The nucleic acids of the
CC invention are useful for assaying the presence of a nucleic acid molecule
CC and for chromosome mapping. They are also used in gene therapy and
CC antisense therapy. The NARC sequences are useful for treating central
CC nervous system disorders and disorders involving aberrant apoptosis, for
CC inducing an immune response and for isolating binding partners. Diseases
CC treated include spleen disorders (e.g. tuberculosis and congestive
CC splenomegaly), lung disorders (e.g. adult respiratory distress syndrome,
CC Goodpasture's syndrome and bronchial asthma), liver disorders (e.g.
CC jaundice and hepatic failure), infectious disorders (e.g. viral
CC hepatitis), brain disorders (e.g. cerebral oedema, hypertensive
CC encephalopathy and hydrocephalus), heart disorders (e.g. heart failure,
CC angina and myocardial infarction), kidney disorders (e.g. cysts and
CC glomerulonephritis), testes and epididymis disorders (e.g. gonorrhoea and
CC syphilis), skeletal muscle disorders (e.g. tumours) and pancreatic
CC disorders (e.g. pancreatitis and diabetes). The present sequence is human
CC neuronal apoptosis regulated candidate (NARC) 16B DNA
XX
SQ Sequence 3206 BP; 988 A; 580 C; 682 G; 956 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 3206
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-10-047-855-3 (1-672) x AAD06010 (1-3206)
Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 145 ATGACACCTTCTCAGGTGGCTTTGAAATAGAGGAACTCTTTTACGAGGAGTTT 204
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTyrAsnProGlnAsnAlaValAlaLeu 40
Db 205 GCGATATGTGGAGCTGTGATGCTTTGGAAACTGGAATCTCTCAAAATGCTGGGCTCT 264
Qy 41 LeuProGluAsnAspThrGlyGluSerMetLeuTyrLysAlaThrIleValLeuSerArg 60
Db 265 CTTCCAGAGAAATGACACAGGTGAAAGCATGCTATGAAAGCAACCATTTGCTCAGTAGA 324
Qy 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 325 GGAGTATCAGTTTCACTGCTACTTCAAGGGGTACTTTTGAACCAAGACTATCGGT 384
Qy 81 GlyProCysGlnValIleValHisLysTyrGluThrHisLeuGlnProArgSerIleThr 100
Db 385 GGTCCATGTCAAGTATAGTTCACAGTGGGAGACTCATCTACACCAACGATCAATACC 444
Qy 101 ProLeuGluSerGluIleIleLeuAspAspGlyGlnPheGlyIleHisGlnValGlu 120
Db 445 CCTTTAGAAAGCGAAATATTATTATGACGATGCAATTTGGAATCCACAATGGTGTGAA 504
Qy 121 ThrLeuAspSerGlyTyrLeuThrCysGlnThrGluIleArgLeuGluHisTyrSer 140
Db 505 ACTCTGGATTTCTGGATGGCTGACATGTCTAGATGAAATTAAGATTACGTTTGCATTATCT 564
Qy 141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
Db 565 GAAAAACCTCTCTGTGTCATAACCAAGAAAAAATTTAAAAAATCTAGATTAGGTGAAG 624
Qy 161 LeuThrLeuGluGlyLeuGluAspAspAspArgValSerProThrValLeuHis 180

625	CtGACACTAGAGCCCTGGAGGAATGACGATGATAGCGTATCTCCCACTGTACTCCAC	684
181	LyeMetSerAsnSerLeuGlulSerLeulleleSerAspAsnGluPheLysCysArgHis	200
685	AAAATGTCAAATAGCTTGAGATATCTTAATAAGCGCAATAGAGTTCAAAGTGCAGGAT	744
201	SerGlnProGluCysGlyTyrglyLeuGlnProAspAtgTrpThrGluTy-SerileGln	220
745	TCACAGCCGAGTGTGGTATATGGCTTGGACGCTGATCGTTGGACAGAGTACAGCATACAG	804
221	ThrMetGluProAspAsnLeuGluLeullePheAspPheGluGluAspLeuSerGlu	240
805	ACGATGGAACAGATTAACCTGGAACTAATCTTTGATTCTTTTCGAAGAAGATCTCAGTGAG	864
241	HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer	260
865	CACGTAGTTACGGGTGATGCCCTCTCTGGACATGTGGGTACAGCTTGTCTTATCATCCC	924
261	ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn	280
925	ACCATTGTCAGAGTGGAAAGAGTGTGGAAATCTTACTCTTCCCCATCATGACGAANAAT	984
281	SerArgLysThrIleGlyLysValatqValaspTyrllellelellysProLeuProGly	300
985	TCCCGGAAAAACAATAGGCAAGTGAGAGTTGACTATATTAATTAATTAAGCCATTACACGA	1044
301	TyrSerCysAspMetLysSerSerPheSerLysTyrrTpLysProArgIleProLeuAsp	320
1045	TACAGTTGTGACATGNAATCTTCATTTCCAAGTATTGGAAAGCAAGAAATACCATTTGGAT	1104
321	ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlalysValGlnGlu	340
1105	GTTGGCCATCGAGGTGCAGGAACTCTACAACTGCCCCAGCTGGCTAAAGTTCAAGAA	1164
341	IanThrIleAlaSerLeuAtqAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp	360
1165	AATACTATTTGCTCTTTAAGAAATGCTGCTAGTAGTATGTCAGCCCTTTGTAGNAITTGAC	1224
361	ValHisLeuSerLysAspPheValproValValTyrrHisAspLeuThrCysCysLeuThr	380
1225	GTACACCTTTCAAGAGACTTTGTGCCGTGTATATCATGATCTTACTCTGTTGTAGCT	1284
381	MetLysLysPheAspAlaAspProValGluLeuPheGluLeuProValLysGluLeu	400
1285	ATGAAAAAGAAATTTGATGCTGATCCAGTTGAATATTGAAATTCAGGTAAAGAAATTA	1344
401	ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp	420
1345	ACATTTGACCAACTCCAGTTGTAAAGCTCACTCATGTGACTGCATCGAATCTTAGGAT	1404
421	ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer	440
1405	CGAAAGAAATCTGTGTTGAGGAGGAAATTCCTTTTCAGAAATCAGCCATTTCTCTCT	1464
441	LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluLysTrp	460
1465	CTTAAGATGTTTTAGAGTCTTTGCGCAGAGATGTAGGTTTAAATTAAGAAATAAATGG	1524
461	IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrrPheAspMetAsn	480
1525	ATCTGCCAGCAAGGATGGATGTGGATGTAACTTATCAACATATTTTGAATGAAT	1584
481	LeuPheLeuAspIlelleleLysThrValLeuGluAsnSerGlyLysArgArgIleVal	500
1585	CTGTTTTTGGATATAATTTTAAAACTGTTTTAGAAATTTCTGGAGAGGAGATAGTG	1644
501	PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrrPro	520
1645	TTTTCTTCATTTGATGCAGATATTTGCAACATGTTTGGCAAAAGCAGAACAAATATCCG	1704
521	IleLeuPheLeuThrGlnGlyLysSerGluIleTyrrProGluLeuMetAspLeuArgSer	540

1705 ATACTATTTTTAACTCAAGGAAAATCTGAGATTTATCTCTGACTCATGGACCTCAGATCT 1761

541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal 560

1765 CGGACAAACCCCAATTTGCAATGAGCTTTTGCACAGTTTGAATAATCTACTGGGATTAATGTA 1824

561 HisThrGluAspLeuLeuAtqAsnProSerTyrrileGlnGluAlalysAlalysGlyLeu 580

1825 CATACTGAAGACTTTGCTCAGAAACCCATCTCTATATTCACAGGCGMAAGCTAAGGACTA 1884

581 ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600

1885 GTCATATTTCTGCTGGGTGATGATACCAATGATCTCTGAAAACAGAGAAATTTGAAGAA 1944

601 LeuGlyValAsnGlyLeulleTyrrAspArgIleTyrrAspTrpMetProGluGlnProAsn 620

1945 CTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATTTGGATGCTCACCAACCAANAT 2004

621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640

2005 ATATTTCCAAGTCGAGCAATTTGGAACGCTGAAGCAGGAAATTCACAGAGCTTAGAGCTGT 2064

641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660

2065 TTGTGTCCACCTGTAGCGCTTTGTTCCCTCACTTTTGTGTGGGAGTCTGATATCCAT 2124

661 ValAlaAsnGlyIleAspAsnValGluAsnAla 672

2125 GTGGATGCCAAGCGCATTTGATAACGTGGAGAAATGCT 2160

RESULT 3

ABS56723 standard; cDNA; 3206 BP.

XX ABS56723;

XX 24-FEB-2003 (first entry)

XX Human NARC16 cDNA.

NARC10; NARC16; cytostatic; immunosuppressive; dermatological; cardiant;
antiinflammatory; nephrotropic; anti-HIV; nootropic; neuroprotective;
antianemic; cerebroprotective; vasotropic; antidiabetic; anticonvulsant;
immunosuppressive; thyromimetic; immunostimulant; antimanic; hypotensive;
tranquilliser; neurologic; gene therapy; gene mapping; apoptosis;
AIDS; cell cycle disruption; programmed cell death regulation;
viral infection; nucleosome assembly; phosphate homeostasis;
cell cycle regulation; cancer; follicular lymphoma; carcinoma;
p53 mutation; graft rejection; hormone-dependent tumour;
autoimmune disorder; valvular heart disease;
systemic lupus erythematosus; diabetes; virus-induced lymphocyte depletion;
immune-mediated immunodeficiency syndrome; neurodegenerative disease; stroke;
Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
spinal muscular atrophy; retinitis pigmentosa; cerebellar degeneration;
myelodysplastic syndrome; ischaemic injury; myocardial infarction;
reperfusion injury; liver disease; idiopathic dilated cardiomyopathy;
ischaemic cardiomyopathy; aplastic anaemia; chronic neutropenia; mania;
myelodysplastic syndrome; central nervous system disorder; anxiety;
senile dementia; Huntington's disease; hypertension; schizophrenia;
severe bipolar affective disorder; human; chromosome 4ql1-4q21; ss.

Homo sapiens.

XX Key Location/Qualifiers
FH 145..2163
FT /*tag= a
FT /product= "NARC16"

XX WO200281516-A2.
XX PD 17-OCT-2002.
XX

PF 16-JAN-2002; 2002WO-US001098.
XX
XX 16-JAN-2001; 2001US-0262306F.
PR 15-JAN-2002; 2002US-00047855.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX Chiang LW;
PI
XX WPI; 2003-058503/05.
DR P-PSDB; ABB84606.
XX
XX Novel isolated programmed cell death-related polypeptide, NARC10 and
PT NARC16, useful for treating disorders associated with abnormal apoptotic
PT process e.g. Alzheimer's disease, cancer, myocardial infarction, stroke.
XX
XX Claim 1; Fig 4A-C; 123pp; English.
XX
XX This invention describes novel cell death-related polypeptides NARC10 and
CC NARC16, located on chromosome 4q11-4q21 and which have cardiant, antiHIV,
CC immunosuppressive, dermatological, antiinflammatory, cerebroprotective,
CC neurotropic, neuroprotective, antianaemic, cardiant, vasotropic, antimanic,
CC antidiabetic, immunosuppressive, cytostatic; thyromimetic, nephrotropic,
CC immunostimulant, anticonvulsant, tranquiliser, hypotensive and
CC neuroleptic activity and can be used in gene therapy. The products of the
CC invention can be used to modulate NARC10 or NARC16 polypeptides or
CC polynucleotides, to map NARC genes on a chromosome, e.g. to locate gene
CC regions associated with genetic disease or to associate NARC10 or NARC16
CC with a disease. The polypeptides are also useful for modulating the
CC apoptotic process, and are therefore useful for modulating, and treating
CC disorders associated with increased apoptosis, inhibition of apoptosis or
CC disruptions in cell cycle, for regulating cellular functions including
CC programmed cell death, nucleosome assembly, phosphate homeostasis and the
CC cell cycle. Preferably, the products of the invention are useful for
CC treating disorders associated with abnormally low rate or abnormally high
CC rate of apoptosis e.g. cancers including follicular lymphomas, carcinomas
CC with p53 mutations, or hormone-dependent tumours, autoimmune disorders
CC including systemic lupus erythematosus, diabetes, graft rejection,
CC Hashimoto's thyroiditis and immune-mediated glomerulonephritis and viral
CC infections e.g. infections caused by herpes viruses, virus-induced
CC lymphocyte depletion (including acquired immunodeficiency syndrome
CC (AIDS)), neurodegenerative diseases manifested by loss of specific sets
CC of neurons (including Alzheimer's disease, Parkinson's disease,
CC amyotrophic lateral sclerosis, spinal muscular atrophy, retinitis
CC pigmentosa, and cerebellar degeneration), myelodysplastic syndromes
CC (including aplastic anaemia), ischaemic injuries (including myocardial
CC infarction, stroke and reperfusion injury), and toxin (e.g. alcohol)
CC induced liver disease, idiopathic dilated cardiomyopathy, ischaemic
CC cardiomyopathy and valvular heart disease, aplastic anaemia, chronic
CC neutropenia, and myelodysplastic syndromes, central nervous system
CC disorders, senile dementia, Huntington's disease, hypertension,
CC schizophrenia, attention deficit disorder, mania, anxiety, severe bipolar
CC affective disorder (BP-I). This sequence encodes the human NARC16 protein
CC described in the method of the invention
XX
SQ Sequence 3206 BP; 988 A; 580 C; 682 G; 956 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3206
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-047-855-3 (1-672) x AB956723 (1-3206)

QY 1 MetThrProSerGlnValAlaPheGluLeuArgGlyThrLeuLeuProGlyGluValPhe 20
DB 145 ATGACACCTTCTCAGGTTGCGCTTGAATATAGAGGAACTCTTTTACCAGAGAGTTT 204
QY 21 AlaIleCysGlySerCysPheAlaLeuGlyValAsnTrpAsnProGlnAsnAlaValAlaLeu 40
|||||

Db 205 GCGATATGTGGAGCTGTGATGCTTTGGGAAACTGGAATCCTCAAAATGCTGTGGCTCTT 264
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTyrLysAlaThrIleValLeuSerArg 60
265 CTTCCAGAGAATGACACAGGTGAAAGCATGCTATGGAAGCAACCATTTGCTACTCAGTAGA 324
QY 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
325 GGAGTATCAGTTTCAGTATCGCTACTTCAAAGGGTACTTTTATAGAACCAAGACTATCGGT 384
QY 81 GlyProCysGlnValIleValHisLysTyrGluThrHisLeuGlnProArgSerIleThr 100
385 GGTCCATGTCAAGTATGATGTTCAAGTGGGAGACTCATCTACACCAGCATCAATAACC 444
QY 101 ProLeuGluSerGluIleIleLeuAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
445 CCTTTAGAAAGCGAAATATTATTATTCACGATGACAAATTTGGAATCCACAATGGTGTGAA 504
QY 121 ThrLeuAspSerGlyTyrPleuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
505 ACTCTGGATTCTGGATGGCTGACATGTGACACTGGAATTAAGATTACGTTTGCAATTATCT 564
QY 141 GlnLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
565 GAAAAACCTCTCTGTGTCAATAACCAAGAAATAATTAATAAATCTAGATTAGGGTGAAG 624
QY 161 LeuThrLeuGluGlyLeuGluAspAspAspArgValSerProThrValLeuHis 180
625 CTGACACTAGAGGGCTGGAGGAAGATGACGATAGAGGTATCTCCCACTGACTCTCCAC 684
QY 181 LysMetSerAsnSerLeuGluLeuSerLeuIleSerAspAsnGluPheLysCysArgHis 200
685 AAAATGTCCAATAGCTTGGAGATATCTTTAATAGCGACAATGAGTCAAGTCAGGCAT 744
QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTyrTrpGluTyrSerIleGln 220
745 TCACAGCCGAGTGGTGTATGGCTTGCAGCTGATCGTGGACAGAGTACAGATACAG 804
QY 221 ThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGlu 240
805 ACGATGGAACACAGATAAATCTGGAATCTTTTGAATTTTTCGAAGAAGATCTCAGTGAG 864
QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
865 CACGTAGTTTCAGGGTGCCTCTCTGGACATGTGGGTACAGCTGTCTCTTATCATCC 924
QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
925 ACCATTGTGAGAGTGGAAAGAGTGTGGAATCTTTACTCTTCCCATCATGAGCAAAAT 984
QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGly 300
985 TCCCGGAAACAAATAGGCAAGTGAAGTGTGACTATATAATTAATTAAGCCATTACCAAGA 1044
QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
1045 TACAGTTGTGACATGAAATCTTCAATTTCCAAAGTATTGAAGCCCAAGAAATACCATGGAT 1104
QY 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
1105 GTTGGCCATCGAGGTGAGGAACCTCAACAACATGCCCGCTGGCTTAAGTTCAAGAA 1164
QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyValAlaPheValGluPheAsp 360
1165 AATACTATTGCTTCTTTAAGAAATGCTGCTAGTATGCTGTCAGCCCTTGTAGAAATTTGAC 1224
QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
1225 GTACACCTTTCAAGAGCATTTGTGCCCGGTGATATCATGATCTTACCTGTTGTTGACT 1284
QY 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
1285 ATGAAAAAGAAATTTGATGCTGATCCAGTTGAATTAATTTGAAATTCAGTAAGAAATTA 1344

cerebroprotective; dermatological; virucide; neuroprotective; phosphatidylglycerolphosphate synthase; GPG synthase; human; ss.

Qy 401 ThrPheAspGlnLeuGlnLeuLeuLeuLeuLeuHisValThrAlaLeuLysSerLysAsp 420
Db 1345 ACATTGACCACTCCAGTGTGTTAAAGCTCCTCATGTGACTGCACCTGAAATCTAAGGAT 1404
Qy 421 ArgLysGluSerValValGlnGlnGluLysSerPheSerGluAsnGlnProPheProSer 440
Db 1405 CGGAAGAACTCTGTGTTTCAGGAGAAATTCCTTTTCAGAAATTCAGCCATTTCTCTTCT 1464
Qy 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
Db 1465 CTTAAGATGGTTTGTAGAGTCTTTGCCAAGATGTAGGGTTTAACTTGAATTAATAATGG 1524
Qy 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTrpPheAspMetAsn 480
Db 1525 ATCTGCCAGCAAGGATGGATGTGGATGGTAACTTAATCAACATATTTTGACATGAAT 1584
Qy 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal 500
Db 1585 CTGTTTTTGGATATAATTTTAAATACTGTTTGTAGAAATTCCTGGAAAGAGGAAATAGTG 1644
Qy 501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTrpPro 520
Db 1645 TTTTCTTCATTGATGCGAGATATTGCAATGGTTCGGCAAAAGCAGCAATATATCCG 1704
Qy 521 IleLeuPheLeuThrGlnGlyLysSerGluIleTrpProGluLeuMetAspLeuArgSer 540
Db 1705 ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGAACTCATGGACCTCAGATCT 1764
Qy 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal 560
Db 1765 CGGCAACACCCCATTTGCAATGAGCTTTGCACAGTTTGAATACTACTGGGATAAATGTA 1824
Qy 561 HisThrGluAspLeuArgAsnProSerTrpIleGlnGluAlaLysAlaLysGlyLeu 580
Db 1825 CATACTGAAGACTGCTCAGAAACCCATCCTATATTTCAAGAGCGCAAAAGCTTAAGGACTA 1884
Qy 581 ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
Db 1885 GTCATATTCCTGCTGGGTGATGATACCATGATCTGTAACAGAGAAATTTGAAGGAA 1944
Qy 601 LeuGlyValAsnGlyLeuIleTrpAspArgIleTrpAspTrpMetProGluGlnProAsn 620
Db 1945 CTTGGAGTTAAATGCTTAATTTATGATAGATATATGATGGATGCTTGAACCAACAAAT 2004
Qy 621 IlePheGlnValGlnGlnLeuArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
Db 2005 ATATTCCAAAGTGAGCAATTTGGAACCGCTGAAGCAGAAATTTGCCAGAGCTTAAGAGCTGT 2064
Qy 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 2065 TTGTGTCCTACTGTTAGCCGCTTTGTTCCCTCATCTTTGTGGGGAGTCTGATATCCAT 2124
Qy 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2125 GTGGATGCCAAGCGCATTTGATAACGTGGAGATGCT 2160

RESULT 4
ADJ57933
ID ADJ57933 standard; cDNA; 3206 BP.

XX AC ADJ57933;

XX DT 06-MAY-2004 (first entry)

XX DE Human NARC 16B cDNA.

XX KW Neuronal apoptosis regulated candidate; NARC; diagnosis; cancer;
XX KW lung disease; cirrhosis; hepatitis; atherosclerosis;
XX KW myocardial infarction; inflammation; anaemia; glomerulonephritis;
XX KW osteoporosis; AIDS; acquired immunodeficiency syndrome;
XX KW Parkinson's disease; Alzheimer's disease; stroke; dermatitis;
XX KW drug screening; gene therapy; cytostatic; hepatotropic; nootropic;

KW cerebroprotective; dermatological; virucide; neuroprotective;
KW XX phosphatidylglycerolphosphate synthase; GPG synthase; human; ss.
OS XX Homo sapiens.
XX PN US2004009553-A1.
XX PD 15-JAN-2004.
XX PF 30-APR-2003; 2003US-00426776.
XX PR 27-SEP-1999; 99US-00406045.
XX PR 22-OCT-1999; 99US-0161188P.
XX PR 31-JAN-2000; 2000US-00495823.
XX PR 28-FEB-2000; 2000US-0185517P.
XX PR 20-OCT-2000; 2000US-00692785.
XX PR 31-JAN-2001; 2001US-00773426.
XX PR 28-FEB-2001; 2001US-00795691.
XX PR 31-OCT-2001; 2001US-0335003P.
XX PR 25-MAR-2002; 2002US-00105992.
XX PR 28-AUG-2002; 2002US-00229662.
XX PR 30-OCT-2002; 2002US-00284014.
XX PR 30-OCT-2002; 2002US-00284059.
XX PR 09-DEC-2002; 2002US-00314881.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX XX Glucksmann MA, Williamson MJ, Tsai P, Rudolph-Owen LA;
XX PI Kapeller-Libermann R, Meyers RE, Chiang LW, Hunter JJ;
XX XX WPI; 2004-090469/09.
XX XX New isolated nucleic acid molecules and polypeptides (e.g. 27411, 23413,
XX PT 22438, 23553, NARC SCI or NARC 1) useful for diagnosing, preventing or
XX PT treating disorders associated with the protein, e.g. cancer,
XX PT atherosclerosis or AIDS.
XX PS Claim 1; SEQ ID NO 34; 260pp; English.
XX XX The present invention provides isolated nucleic acid molecules and
XX CC proteins designated 27411, 23413, 22438, 23553, 25278, 26212, NARC SCI,
XX CC NARC 10A, NARC 1, NARC 12, NARC 13, NARC 17, NARC 25, NARC 3, NARC 4,
XX CC NARC 7, NARC 8, NARC 11, NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20,
XX CC NARC 26, NARC 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C,
XX CC NARC 8B, NARC 9, NARC 2A, NARC 16B, NARC 1C, NARC 1A, NARC 25, 86604 or
XX CC 32222. The invention is useful in diagnosing, preventing or treating
XX CC disorders such as cancer, lung diseases, cirrhosis, hepatitis,
XX CC atherosclerosis, myocardial infarction, inflammation, anaemia,
XX CC glomerulonephritis, osteoporosis, AIDS (acquired immunodeficiency
XX CC syndrome), Parkinson's disease, Alzheimer's disease, stroke and
XX CC dermatitis. These may also be used in drug screening. The invention is
XX CC also useful in gene therapy. The present sequence is human neuronal
XX CC apoptosis regulated candidate (NARC) cDNA.
XX SQ Sequence 3206 BP; 988 A; 580 C; 682 G; 956 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3206
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-047-855-3 (1-672) x ADJ57933 (1-3206)

Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 145 ATGACACCTTCTCAGTTCCTTTCGAATAAGAGGAACTCTTTTACCAGGAGAGTTTTT 204
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40

DB 205 GCGATATGTGGAAGCTGTGATGCTTTTGGGAAACTGGAATCCTCAAAATGCTGTGGCTCTT 264
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
DB 265 CTTCCAGAGAAATGACACAGCTGAAAGCATGCTATGGAAGCAACCACTTGTACTCACTAGA 324
QY 61 GlyValSerValGlnTrpArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
DB 325 GGAGTATCAAGTTCAGTATCGCTACTCTCAAGGGGTACTTTTGAACCAAGACTATCGGT 384
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
DB 385 GTTCCATGTCAGTATAGTATTCACAAGTGGAGACTCATCTACACCAAGATCAATAACC 444
QY 101 ProLeuGluSerGluIleIleLeuAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
DB 445 CCTTTAGAAAGCGAAATATTATTGACGATGGACAATTTGGAATCCACAATGCTGTGAA 504
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgIleHisTyrSer 140
DB 505 ACTCTGGATTCTGGATGGCTGACATGTCAGACTGAAATAAGATTACGTTTGCATTATTCT 564
QY 141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
DB 565 GAAAAACCTCCTCTGTGTCATAACCAAGAAAAAATTAAAAAATCTAGATTGAGGTGAAG 624
QY 161 LeuThrLeuGluGlyLeuGluAspAspAspArgValSerProThrValLeuHis 180
DB 625 CTGACACTAGAGCCCTGGAGNAGATGACGATGATAGGGTATCTCCCACTGTACTCCAC 684
QY 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
DB 685 AAAATGTCATAGCTTGGAGATATCTTAATAAGCGCAATGAGTTCAAGTCAGGCAT 744
QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpGluTrpSerIleGln 220
DB 745 TCACAGCCGGAGTGTGGTTATGGCTTGCAGCCCTGATCGTTGGACAGACGATACACAG 804
QY 221 ThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGlu 240
DB 805 ACGATGGAAACCATAGTAACCTGGNACTAATCTTTGATTTTTCGAAGAAGATCTCAGTGAG 864
QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
DB 865 CACGTAGTTTCAGGGTCATGCCCTTCTCGACATGTGGTACAGCTTGTCTCTTATCATCC 924
QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrIleuProIleMetSerArgAsn 280
DB 925 ACCATTGCTGAGAGTGGAAAGAGTGTGGAATTTCTTACTCTTCCCAATCATGACAGAAAT 984
QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300
DB 985 TCCCGGAAACAATAGCAAGATGAGAGTTGACTATATAATTATTAAAGCCATTACACGGA 1044
QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
DB 1045 TACAGTTGTGACATGAAATCTTCAATTTTCCAAAGTATGGAAGCCAAAGATACCATTTGGAT 1104
QY 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
DB 1105 GTTGGCCATCGAGGTGCAGAAACTCTACAACAACTGCCAGCTGGCTAAAGTTCAAGAA 1164
QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
DB 1165 AATACTATTGCTTCTTTAGAAATGCTGCTAGTCATGGTGCAGCCCTTTGTAGAAATTGAC 1224
QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
DB 1225 GTACACCTTTCAAAGACCTTTGTGCCCGTGGTATATCATGATCTTACCTGTTGTTGACT 1284
QY 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
DB 1285 ATGAAAAAGAAATTTGATGTCATCCAGTTGAATTTTGAATTTCCAGTAAAGAAATTA 1344

QY 401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
DB 1345 ACATTTGACCAACTCCAGTTGTAAAGCTCACTCATGTGACTGCACATAAATCTAAGGAT 1404
QY 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
DB 1405 CGGAAGAATCTGTGGTTCAGAGGAAATTCCTTTTCAGAAATCAGGCATTTCTCTCT 1464
QY 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
DB 1465 CTTAGATGGTTTGTAGCTCTTTGCCAAGAGATGTAGGGTTTAAACATTGAAATATAATGG 1524
QY 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
DB 1525 ATCTGCCAGCAAAAGGATGGAATGTGGATGGTAACCTTATCAACATATTTTGCATGAAT 1584
QY 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal 500
DB 1585 CTGTTTTTGGATATATATTTTAAAAACTGTTTTAGAAAAATTTCTGGGAAGAGGAAATAGTG 1644
QY 501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro 520
DB 1645 TTTTCTTCAATTTGATGCAGATATTTGCACAAATGGTTCCGCAAAAGCAGCAACAATATCCG 1704
QY 521 IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer 540
DB 1705 ATACTATTTTAACTCAAGGAAAACTGAGATTTATCTCTGAACTCATGCACCTCAGATCT 1764
QY 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal 560
DB 1765 CGGACAAACCCCATTTGCAATGAGCTTTGCACAGTTTGAATAATCTACTGGGGATAAATGTA 1824
QY 561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
DB 1825 CATCTGAAGACTTGTCTCAGAAACCCATCTATATTCAAGAGGCAAAAGCTAAGGACCTA 1884
QY 581 ValIlePheCysTrpGlyAspThrAsnAspProGluAsnArgLysLeuLysGlu 600
DB 1885 GTCATATTTCTGTGGGTGATGATACCAATGATCTCTGAAAAACAGAAAGGAAATGAAGGAA 1944
QY 601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
DB 1945 CTTGAGTTTAAATGGTCTAATTTTATGATAGATATATGATTGGATGCTCGAACCAACCAAT 2004
QY 621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
DB 2005 ATATTCCAAAGTGGAGCAATTGGAAACCCCTGAAGCAGAAATTCAGAGGCTTAAGAGCTGT 2064
QY 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
DB 2065 TTGTGTCCCACTGTTAGCCGCTTTGTTCCCTCATCTTTGTGGGGAGTCTGATATCCAT 2124
QY 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
DB 2125 GTGGATGCCAACGGCATTGTAACGTGGAGAAATGCT 2160
RESULT 5
ABV29614
ID ABV29614 standard; cDNA; 3499 BP.
XX AC ABV29614;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 29605.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN W0200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 6349; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 3499 BP; 1104 A; 609 C; 736 G; 1040 T; 0 U; 10 Other;

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0	3537.00	3499	672
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-10-047-855-3 (1-672) x ABV29614 (1-3499)

QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20

DB 221 ATGACACCTTCTCAGGTGCTTGGAAATAGAGGAACCTTTTACCAGGAGAGTTT 280

QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40

DB 281 GCATATGTGGAAGCTGTGATGCTTTGGGAACTGGAATCCTCAAAATGCTGGCTCT 340

QY 41 LeuProGlnAsnAspThrGlyGluSerMetLeuTrpIleAlaThrIleValLeuSerArg 60

DB 341 CTTCACAGAAATGACACAGGTGAAGCATCTATGGAAGCAACCACTTGTACTCAGTAGA 400

QY 61 GlyValSerValGlnTyrArgTyrPheLeuGlyTyrPheLeuGluProIleThrIleGly 80

DB 401 GGAGTATCAGTTCAGTATCGCTCTTCAAGGGGTACTTTTAGAACCAAGACTACCGT 460

QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100

DB 461 GGTCCATGTCAGGTGATAGTTCACAGTGGGAGACTCATCTACAAACCAACATCAATACC 520

QY 101 ProLeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120

DB 521 CCTTTAGAAAGCGAAATATTATTATGACATGGACATTTGGAAATCCCAATGGTGTGAA 580

QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140

DB 581 ACTCTGGATTCTGGATGCTGACATGTCAGATGAATAAGATTACGTTGTCATTATTCT 640

QY 141 GluLysProProValSerIleThrLysLysLeuLysLysSerArgPheArgValLys 160

DB 641 GAAAAACCTCTCTGTCTCAATAACCAAGAAAAATTAATAATCTAGATTTAGGGTGAG 700

QY 161 LeuThrLeuGluGlyLeuGluGluAspAspArgValSerProThrValLeuHis 180

DB 701 CTGACACTAGAAAGGCTCGAGGAAGATGACGATGATAGGGTATCTCCCACTGTATCCAC 760

QY 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200

DB 761 AAAATGTCCATAGCTTGGAGATATCTTAAAGCGCAATAGTTCAGTGCAGGCAT 820

QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220

DB 821 TCACAGCGGAGTGTGTTATGCTTGCAGCCCTGATCGTTGGACAGAGTACAGCATACAG 880

QY 221 ThrMetGluProAspAsnLeuGluIlePheAspPheGluGluAspLeuSerGlu 240

DB 881 ACGATGGAACCAAGATTAACCTGGAACATACTTTGATTTTTTGAAGAGATCTCAGTGAG 940

QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260

DB 941 CAGTATGTCAGGTGATGCCCTTCTCGACATGTGGTACAGCTGTCTCTTATCATCC 1000

QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280

DB 1001 ACCATTGCTGAGAGTGGAAAGAGTGTGGAATTTCTTACTTCTCCCATCATGACAGAAAT 1060

QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300

DB 1061 TCCCGGAAAACAATAGCAAAAGTGAGATGATATATATATTAATTAAGCCATTACCAGGA 1120

QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320

DB 1121 TACAGTTGTCATGAATCTTTCATTTTCCAGATTTTGAAGCAAGAAATACCATTTGGAT 1180

QY 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340

DB 1181 GTTGCCCATCGAGTGCAGGAAACTCTACAACTCTGCCAGCTGGTAAAGTTCAAGAA 1240

QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360

DB 1241 AATACTATTGCTCTTTAAGAAATGCTGCTGATGTCATGTCAGCTTTGTAGAATTGAC 1300

QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380

DB 1301 GTACACCTTTCAGAGGACTTTGTGCCGTGGTATATCATGATCTTACCTGTGTTGACT 1360

QY 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400

DB 1361 ATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTTGAATTTCCAGTAAAAAGAAATTA 1420

QY 401 ThrPheAspGlnLeuGlnLeuLeuLeuLeuThrHisValThrAlaLeuLysSerLysAsp 420

DB 1421 ACATTGTCCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACCTGAATCTAAGGAT 1480

QY 421 ArgLysGluSerValValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSer 440

DB 1481 CGGAAAGAAATCTGTGTTTCAGGAGAAATTTCTTTTCAGAAATTCAGCCATTCTCTCT 1540

QY 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460

DB 1541 CTTAAGATGTTTATGAGTCTTTGTCAGAAAGATGAGGGTTTAACTTGAATAATAAATGG 1600

QY 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480

DB 1601 ATCTCCCAAGCAAGGATGGAATGTGGATGGTAATCTTATCAACATATTTTTCACATGAAT 1660

QY 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal 500

Db 1661 CTTGTTTGGATATAATTTTAAACACTGTTTAAAGAAATTTCTGGGAAGGAGAAATAGTG 1720
QY PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro 520
Db 1721 TTTTCTTCATTCATGACATATTTGACATATGTTTCGGAAGAGCAGACAAATATCCG 1780
QY IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer 540
Db 1781 ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTCTGAACCTCATGGACCTCAGATCT 1840
QY ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal 560
Db 1841 CGGACAAACCCCAATGCAATGAGCTTTGCACAGTTTGAAAATCTTACTGGGGATAAATGTA 1900
QY HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
Db 1901 CATACTGAAGACTTGTCTCAGAACCCATCTTATATTCAGAGGCAAAAGCTAAGGACTA 1960
QY ValIlePheCysTrpGlyAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
Db 1961 GTCATATTCCTGCTGGGGTGATGATACCAATGATCTCTGAAAACAGAGGAATTTGAAGGA 2020
QY LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
Db 2021 CTTGGAGTTAATGCTCTAATTTATGATAGGATATATGATTGGATGCTGAACCAACCAAT 2080
QY IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
Db 2081 ATATTCAGAGTGAGCAATTTGGACCGCTGACAGCAAGAAATTCAGAGCTTAAGAGCTGT 2140
QY LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 2141 TTGTGTCCCACTGTAGCGCTTTGTTCCTCATCTTTGTGTGGGAGTCTGATATCCAT 2200
QY ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2201 GTGGATGCCAACGGCATTGATAACGTTGGAGAAATGCT 2236

RESULT 6

ID ADL62788 standard; DNA; 3499 BP.

XX ADL62788;

XX 20-MAY-2004 (first entry)

XX Human ovarian cancer DNA marker #21000.

XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

XX Homo sapiens.

XX WO200170979-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009126.

XX 21-MAR-2000; 2000US-0191031P.

XX 25-MAY-2000; 2000US-0207124P.

XX 15-JUN-2000; 2000US-0211940P.

XX 07-JUL-2000; 2000US-0216820P.

XX 25-JUL-2000; 2000US-0220661P.

XX 21-DEC-2000; 2000US-0257672P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lee J, Lillie J;

XX WPI; 2001-611502/70.

XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian

PT cancer cells as compared to their normal non-cancerous ovarian cells are
XX used to characterize stage, grade, histological type of ovarian cancer.
XX Disclosure; SEQ ID NO 21000; 106pp; English.

XX The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.

XX Sequence 3499 BP; 1104 A; 610 C; 736 G; 1039 T; 0 U; 10 Other;

Alignment Scores:

Pred. No.: 0 Length: 3499
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-047-855-3 (1-672) x ADL62788 (1-3499)

QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 221 ATGACACCTTCACAGTTGCTTGGAAATAGAGGAACCTTTTACCGAGAGATTTT 280
QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 281 GCCATATGTGGAAGCTGTGATGCTTTGGGAAACCTGCAAAATCTGTGGCTCTT 340
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 341 CTTCCAGAGAATGACACAGGTGAAGCATGCTATGGAAGCAACCACTTGACTCAGTAGA 400
QY 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 401 GGAGTATCAGTTCACTATCGTACTTCAAGGGGTACTTTTAGAACCAAGACTATCGGT 460
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 461 GGTCCATGTCAAGTGATAGTTTCAAGAGTGGAGACTCATCTACACCAACCAATCAATACC 520
QY 101 ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisGlyValGlu 120
Db 521 CCTTTAGAAAGCGAAATATTATTGACGTGGAGCAATTTGGAAATCCACATGTTGTGAA 580

```
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisIleTyrSer 140
Db 581 ACTCTGGATTCGGATGGCTGACATGTCAGACTGAAATAAGATTACCGTTTGCATTAATCT 640
QY 141 GluLysProValSerIleThrLysLysLysLysLysLysSerArgPheArgValLys 160
Db 641 GAAAAACCTCTGTGTCAATACCAAGAAAAAATTAATAAATAATCTAGATTAGGGTGAAG 700
QY 161 LeuThrLeuGluGlyLeuGluGluAspAspArgValSerProThrValLeuHis 180
Db 701 CTGCACATAGAAAGGCTTCAGAGAAAGATGACGATGATAGGGTATCTCCACATGTAATCCAC 760
QY 181 LysMetSerAsnSerLeuGluIleSerLeuLysSerAspGluPheLysCysArgHis 200
Db 761 AAATGTCCATAGCTTCGAGATATCTTTAATAAGCGACATGATGTTCAAGTGCAGGCAT 820
QY 201 SerGlnProGluCysGlyTrpGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
Db 821 TCACAGCGGAGTGTGGTTATGGCTTGCGAGCTGTATCGTTGGACAGAGTACAGCATACAG 880
QY 221 ThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGlu 240
Db 881 ACGATGGAAACCAAGATAACCTGGAACATACTTTGTGATTTTTCGAAGAAGATCTCAGTGAG 940
QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db 941 CACGTAGTTCCAGGTGATGCCCTTCTCGACATGTGGGTACAGCTTGTCTCTTATCATCC 1000
QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
Db 1001 ACCATTGCTGAGATGGAAAGATGCTGGAAATCTTACTCTTCCCATCATGAGCAGAAT 1060
QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300
Db 1061 TCCCGAAACCAATAGCGAAAGTGGAGTTGATATATTAATTAATTAAGCCATTACCGA 1120
QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
Db 1121 TACAGTTGTGACATGAATCTTCATTTTCCAGTATTGGAAGCAAGATACCATTTGGAT 1180
QY 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGluLeuAlaLysValGlnGlu 340
Db 1181 GTTGGCCATCGAGTGCAGAAACTCTACAAACACTGCCAGCTGGCTAAAGTTCAAGAA 1240
QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyValAlaPheValGluPheAsp 360
Db 1241 AATACTATTGCTCTTTAAGAAATGCTGTAGTCACTGTTGGTGCAGCTTTGTAGAAATTTGAC 1300
QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
Db 1301 GTACACCTTTCAAAGGACTTTGTGCGCGTGGTATATCATGATCTTACCTGTTGTTTGACT 1360
QY 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
Db 1361 ATGAAAAAGAAATTCGATGCTGATCCAGTTGAATTAATTTGAAATTCAGTAAAGAAATTA 1420
QY 401 ThrPheAspGlnLeuGlnLeuLeuLeuLeuLeuLeuHisValThrAlaLeuLysSerLysAsp 420
Db 1421 ACATTTGACCAACTCCAGTTGTTAAGCTCACTCATGTGCTGACATGAAATCTAAGGAT 1480
QY 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
Db 1481 CGGAAAGAAATCTGTGGTTTCAGGAGAAAAATTCCTTTTCAGAAAAATCAGCCATTTCTTCT 1540
QY 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
Db 1541 CTTAAGATGGTTTAGAGTCTTTCCGAGAAGATGTAGGGTTAATCATGTAAGAAATAAATGG 1600
QY 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
Db 1601 ATCTGCCAGCAAGGATGGAAATGTGGATGGTAACCTATCAACATATTTGACATGAAT 1660
QY 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal 500
1661 CTGTTTTTGGATATAATTTTAAAAAACTGTTTAAAAAATTTCTGGGAAGAGGAGAAATAGTG 1720
501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTrpPro 520
1721 TTTTCTTCAATTTGATGATGATATTTGCAACAATGGTTGGCAAAAGCAGAAACAATATCCG 1780
521 IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer 540
1781 ATACTATTTTAACTACAGAAAAATCTGAGATTTATCTCTGAACTCATGACCTCAGATCT 1840
541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
1841 CGACAAACCCCATTCGATGAGCTTTGCACAGTTTAAAAATCTACTGGGGATAAATGTA 1900
561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
1901 CATACTGAAGACTTGTCTCGAAACCCATCTCTATATTCAGAGGCAAAAGCTAAGGACTA 1960
581 ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
1961 GTCATATTTCTGCTGGGTGATGATACCAATGATCTCTGAAACAGAAAGAAATTTGAAGGAA 2020
601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
2021 CTTGGAGTTAATGGTCTAAATTTATGATAGGATATATGATTTGGATGCTGAAACCAACAAAT 2080
621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
2081 ATATTCCAAGTGGAGCAATTTGGAAACGCTTGAAGCAGAAATTTGCCAGAGCTTAAGAGCTGT 2140
641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
2141 TTGTGTCCTTCTGTTAGCCGCTTTGTTCCCTCATCTTTGTGTGGGAGTCTGATATCAT 2200
661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
2201 GTGGATGCCAACGCAATTGATAACGTGGAGAAATGCT 2236
RESULT 7
ADN05171
ID ADN05171 standard; cDNA; 5443 BP.
XX
AC ADN05171;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #801.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
DR
DR WPI: 2004-305105/28.
DR P-PSDB; ADN05172.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
```

PS Claim 1; SEQ ID NO 1565; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for treating psoriasis or a sequence having at least 80% identity to the CC above sequences. The nucleic acid is useful for preparing a composition for diagnosing or treating psoriasis in a mammal. This sequence CC corresponds to one of the polynucleotides of the invention.

XX SQ Sequence 5443 BP; 1544 A; 1034 C; 1149 G; 1716 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	5443
Score:	3537.00	Matches:	672
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-047-855-3 (1-672) x ADN05171 (1-5443)

QY	1	MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe	20
DB	205	ATGACACCTTCTCAGGTGGCTTTGAAATAAGAGGAACTCTTTTACCAGGAGATTTT	264
QY	21	AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu	40
DB	265	GCGATATGTGGAGCTGTGATGCTTTGGGAAACTGGAATCCTCAAAATGCTGTGGCTCT	324
QY	41	LeuProGluAsnAspThrGlyGluSerMetLeuTrpIleValThrIleValLeuSerArg	60
DB	325	CTTCCAGAGAAATGACACAGGTGAAGCATGCTATGGAAGCAACCATTTGACTCAGTAGA	384
QY	61	GlyValSerValGlnTrpArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly	80
DB	385	GGAGTATCAGTTCAGTATCGTACTTCAAGGGTACTTTTAGAACCAAGACTATCGGT	444
QY	81	GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr	100
DB	445	GGTCCATGTCAAGTGTAGTTCACAAAGTGGGAGCTCATCTACACCAACGATCAATAACC	504
QY	101	ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisGlnValGlu	120
DB	505	CCTTTGAAGAGCAAAATATTATTGACGATGGACAATTTGGAATCCACATGTTGTGA	564
QY	121	ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer	140
DB	565	ACTCTGATCTCGATGGCTGACATGTCAGACTGAAATAGATTAGTTTGCATATTCT	624
QY	141	GluLysProValSerIleThrLysLysLysLeuLysSerArgPheArgValLys	160
DB	625	GAATAACCTCTCTGTGTCATTAACCAAGAAAAAATTAATAAATCTAGATTAGGGTGAAG	684
QY	161	LeuThrLeuGluGlyLeuGluAspAspAspAspArgValSerProThrValLeuHis	180
DB	685	CTGACACTAGAGGCTGGAGGAAGATGACGATAGTAGGGTATCTCCCACTGTACTCCAC	744
QY	181	LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis	200
DB	745	AAATGTCCATAGCTTGGAGATATCTTATATAGGCACATGATGATTCAGTGACGGCAT	804
QY	201	SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln	220
DB	805	TCACAGCGGAGTGTGGTTATGGCTTGCAGCTGATCGTTGGACAGATACATACATACAG	864
QY	221	ThrMetGluProAspAsnLeuGluIlePheAspPheGluGluAspLeuSerGlu	240
DB	865	ACGATGGAAACCATTAACCTGGAACTAATCTTTGATTTTTCGAAGAAGATCTCAGTGAG	924
QY	241	HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer	260
DB	925	CAGTAGTTCAGGGTATGCCCTCTCTGGACATGTGGTACAGCTTGTCTTATCATCC	984
QY	261	ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn	280

DB	985	ACCATTTGCTGAGAGTGGAAAGAGTGTGGAAATCTTTACTCTTCCCATCATGACGAAAT	1044
QY	281	SerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGly	300
DB	1045	TCCCGGAAACAATAGGCAAGAGTGAGAGTTGACTATATATATATATATATATACAGGA	1104
QY	301	TyrSerCysAspMetLysSerPheSerLysTyrTrpLysProArgIleProLeuAsp	320
DB	1105	TACAGTTGTGACATGAAATCTTTTCATTTTCAAGTATTTGGAAGCCAAGATACCATTTGAT	1164
QY	321	ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu	340
DB	1165	GTTGGCCATCGAGGTGCAGGAACTCTACAAACACTGCCAGCTGGCTAAAGTTCAAGAA	1224
QY	341	AsnThrIleAlaSerLeuArgAsnAlaLaSerHisGlyValAlaPheValGluPheAsp	360
DB	1225	AATACTATTGCTTCTTTAAGAAATGCTGCTAGTCATGTTGTCAGCTTTGTAGAAATTGAC	1284
QY	361	ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr	380
DB	1285	GTACACCTTTCAAAGGACTTTGTGCCGCTGGTATATCATGATCTTACCTGTTGTTGACT	1344
QY	381	MetLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu	400
DB	1345	ATGAAAGAGAAATTTGATGCTGATCCAGTTGAATATTATTTGAAATTCAGTAAAGAAATTA	1404
QY	401	ThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp	420
DB	1405	ACATTTGACCACTCCAGTTGTTAAAGCTCCTCATGTGCTGCTGCACTGAATCTAAGGAT	1464
QY	421	ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer	440
DB	1465	CGEAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTTTCAGAAATCAGCACTTCTCTCT	1524
QY	441	LeuLysMetValLeuGlnSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp	460
DB	1525	CTTAAGATGGTTTATAGAGTCTTTTGCCAGAGATGTAGGGTTTAAACATTTGAAATG	1584
QY	461	IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn	480
DB	1585	ATCTGCCACAAAGGATGGATGTGGATGTGTAATCTATCAACATATTTTTGCATGAAT	1644
QY	481	LeuPheLeuAspIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal	500
DB	1645	CTGTTTTGATATATATTTTAAACACTGTTTTAGAAAAATTCGGGAAGAGGAGAAATG	1704
QY	501	PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro	520
DB	1705	TTTTTCTTCATTTTATGTCAGATATTTGACAAATGGTTTCGGCAAAAGCAGCAAAATATCCG	1764
QY	521	IleLeuPheLeuThrGlnGlnLysSerGluIleTyrProGluLeuMetAspLeuArgSer	540
DB	1765	ATACTATTTTAACTCAGGAAATCTGAGATTTATCTGAACTCATGACCTCATGATCT	1824
QY	541	ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal	560
DB	1825	CGGACAAACCCCATTTGCAATGAGCTTTGCACAGTTTGAATAATCTACTGGGATAAATGTA	1884
QY	561	HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu	580
DB	1885	CATACTGAAAGATTTGCTCAGAAACCCATCTTATTTCAAGAGGCAAAAGCTAAGGACTA	1944
QY	581	ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu	600
DB	1945	GTCATATTTCTGCTGGGGTATGATACCAATGATCTCTGAAACAGAAAGGAAATGAAAGGA	2004
QY	601	LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn	620
DB	2005	CTTGGATTAATGGTCTAATTTATGATAGATATATGATTTGGATGCTCTGAACCAACAAAT	2064
QY	621	IlePheGlnValGluGlnLeuGluArgLysGlnLeuProGluLeuLysSerCys	640

Db 2065 ATATTCCAGTGGAGCAATTGGAAACGCTGAGCAGGAATTGCAGAGCTTAAGAGCTGT 2124
Qy 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 2125 TTGTGTCCACTGTAGCGGTTTGTTCCTTCATCTTTGTGTGGGAGTCTGATATCCAT 2184
Qy 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2185 GTGGATGCCAACGGCAATTGATAACGTGGAGAACTGCT 2220

RESULT 8
ID ADS74320
ID ADS74320 standard; cDNA; 5443 BP.
AC ADS74320;
XX 16-DEC-2004 (first entry)
XX PRO83903 cDNA clone DNA327983, role in immune-related disease.
XX PRO83903; Crohn's disease; rheumatoid arthritis; gastrointestinal-gen.;
KW antirheumatic; antiarthritic; psoriasis; antipsoriatic; gene; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FH 190..2223
FT /tag= a
FT /product= "PRO83903"
FT /partial
FT /note= "No start codon"

XX WO2004081199-A2.
XX 23-SEP-2004.
XX 10-MAR-2004; 2004WO-US007862.
XX 11-MAR-2003; 2003US-0454025P.
XX (GETH) GENENTECH INC.

XX Baldwin D, Bodary S, Clark H, Fong S, Gurney AL, Williams PM;
XX WPI; 2004-668955/65.
XX P-PSDB; ADS74321.

XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
XX treating psoriasis, Crohn's disease, Ulcerative Colitis, or rheumatoid
XX arthritis.

XX Claim 1; SEQ ID NO 39; 166pp; English.

XX The present sequence is of cDNA clone DNA327983 encoding novel human PRO
XX polypeptide PRO83903. The invention provides newly identified and
XX isolated nucleotide sequences encoding polypeptides referred to as PRO
XX polypeptides that are useful in the diagnosis and treatment of immune-
XX related diseases. Microarray analysis showed that DNA327983 is up-
XX regulated 1.5-fold in lesional skin as compared to non-lesional skin from
XX psoriasis patients, up-regulated 1.3-fold in colon samples from Crohn's
XX disease patients as compared to normal colon and up-regulated 1.3-fold in
XX white blood cells from rheumatoid arthritis patients as compared to those
XX from healthy donors. It is also down-regulated 1.5-fold upon activation
XX of CD4 T cells with CD28 or ICAM and down-regulated 2-fold upon
XX activation of monocytes with LPS. PRO83903 can be used in a claimed
XX method of identifying a compound that inhibits expression of the gene
XX encoding it. The candidate compound is especially an antisense nucleic
XX acid. The PRO polypeptide, its antagonist or an antibody that binds the
XX polypeptide are used in claimed methods for the alleviation or diagnosis
XX of rheumatoid arthritis, Crohn's disease and psoriasis. A vector
XX comprising the present nucleic acid can be used to transform a host cell,
XX especially a CHO cell, Escherichia coli or yeast, for production of the
XX PRO polypeptide.

XX Sequence 5443 BP; 1544 A; 1034 C; 1149 G; 1716 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 0 Length: 5443
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-047-855-3 (1-672) x ADS74320 (1-5443)
Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 205 ATGACACCTTCTCAGGTGGCTTTTGAATAAGAGGAACCTTTTACCAGGAGAAGTTTTT 264
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 265 GCATATGTGAAGCTGTGATGCTTTGGGAAACCTGGAATCCTCAAATGCTGTGGCTCTT 324
Qy 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpIlyAlaThrIleValLeuSerArg 60
Db 325 CTTCCAGAGAATGACACAGGTGAAGCATGCTATGGAAACCAACCATTTGTACTCAGTAGA 384
Qy 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProIlyThrIleGly 80
Db 385 GGAGTATCAGTTCAGTATCCCTACTCTCNAAGGGTACTTTTATAGAACCAAGACTATCGGT 444
Qy 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 445 GGTCCATGTCAAGTGATAGTTCAAGTGGGAGACTCATCTCAACACCAACGATCAATAAAC 504
Qy 101 ProLeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
Db 505 CCTTTAGAAAGCGAAATTTATTTACGATGGGCAATTTGGAATCCCAATGGTGTGAA 564
Qy 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
Db 565 ACTCTGGATTCTGGATGGCTGACATGTCAGACTGAATAGATTACGTTTGCATTATCT 624
Qy 141 GluLysProProValSerIleThrLysLysLysLysLysSerArgPheArgValLys 160
Db 625 GAAAAACCTCTGTGTCATTAACCAAGAAAAAATTAATAAGCAATAGATTAGGTTAGGTTGAG 684
Qy 161 LeuThrLeuGluGlyLeuGluGluAspAspArgValSerProThrValLeuHis 180
Db 685 CTGACACTAGAAGGCTGGAGGAAGATGACATGATAGGGTATCTCCACTGTACTCCAC 744
Qy 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
Db 745 AAAATGTCCAATAGCTTGAGATATCTTTAATAGCGCAATAGATTCAAGTGCAGGCAT 804
Qy 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
Db 805 TCACAGCCGAGTGTGGTTATGGCTTGCAGCTGATCGTTGGACAGAGTACAGCATACAG 864
Qy 221 ThrMetGluProAspAsnLeuGluLeuIlePheAspPhePheGluGluAspLeuSerGlu 240
Db 865 ACGATGGAAACCAAGATAACCTGGAACTAATCTTTTGAATTTTTCGAAAGAAGATCTCAGTGAG 924
Qy 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db 925 CAGGTAGTTACAGGTGATGACCTTCTCTGGACATGTGGGTACAGCTTGTCTTATCATCC 984
Qy 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgan 280
Db 985 ACCATTGCTGAGAGTGGAAAGAGTGTCTGGAATTTCTTACTCTTCCCATCATGAGCAAAAT 1044
Qy 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGly 300
Db 1045 TCCCGGAAAAACAATAGGCAAGTGAGAGTTGACTATATATTTATTATTAAAGCCATTACCAGGA 1104

```
Qy 301 TyrSerCysAspMetIysSerSerPheSerLysTyTrrIysProArgIlePProLeuAsp 320
Db 1105 TACAGTTGTGACATGAATCTTCATTTTCCAGTATTTGGAAGCCAAAGAAATACCATTTGGAT 1164
Qy 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
Db 1165 GTTGGCCATCGAGTGCAGGAACTCTACAACTGCCCGAGCTGGCTAAAGTTCAAGAA 1224
Qy 341 AsnThrIleAlaSerLeuArgAsnAlaIleSerHisGlyAlaAlaPheValGluPheAsp 360
Db 1225 AATACATTATGCTCTTTTAAAGAAATGCTGTAGTCATGCTGAGCCCTTTGTAGAAATTTGAC 1284
Qy 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
Db 1285 GTACACCTTTCAAGAGCTTTGTGCCGCTGTATATCATGATCTTACCTGTGTTGTGACT 1344
Qy 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
Db 1345 ATGAAAAGAAATTTGATGCTGATCCAGTTGAATTTATTTGAAATTCAGTAAAGAATTA 1404
Qy 401 ThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
Db 1405 ACATTTTGACCACTCCAGTTGTTTAAAGCTCCTCATGCTGACTGCACCTGAAATCTAAGGAT 1464
Qy 421 ArgLysGluSerValValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSer 440
Db 1465 CGAAAGAAATCTGTGTTTTCAGGAGGAAATTCCTTTTCAGAAATCAGCCATTTCTCTCT 1524
Qy 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrr 460
Db 1525 CTTAAGATGTTTGAAGCTTTTGCAGAAAGATGTAGGGTTTAACTTGAATTAATAATGG 1584
Qy 461 IleCysGlnGlnArgAspGlyMetTrrAspGlyAsnLeuSerThrTrrPheAspMetAsn 480
Db 1585 ATCTGCCAGCAAGGATGGAATGTGGATGTTAACTTATCAACATATTTTGACATGAAT 1644
Qy 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal 500
Db 1645 CTGTTTTTGGATATAATTTTAAAACTGTTTTAGAAAAATTCGGGAAGAGGAGATAGTG 1704
Qy 501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTrrPro 520
Db 1705 TTTTCTTCATTTGATCGATATTTGCACATGTTTGGCAATGGTTGGCAAAAGCAGCAAAATATCCG 1764
Qy 521 IleLeuPheLeuThrGlnGlyLysSerGluIleTrrProGluLeuMetAspLeuArgSer 540
Db 1765 ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGAACTCATGACCTCATGATCT 1824
Qy 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
Db 1825 CGGACAAACCCCATTCGAATGAGCTTTGCACAGCTTTGMAAATCTACTGGGGATAAATGTA 1884
Qy 561 HisThrGluAspLeuLeuArgAsnProSerTrrIleGlnGluAlaLysAlaLysGlyLeu 580
Db 1885 CATACTGAAGACTGTCTGAGAAACCCATCTATTTCAAGAGCCAAAGCTAAGGACTA 1944
Qy 581 ValIlePheCysTrrPglyAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
Db 1945 GTCATATTTCTGCTGGGTGATGATACATGATCTCTGAAAACAGAGAAATTAAGAGAA 2004
Qy 601 LeuGlyValAsnGlyLeuIleTrrAspArgIleTrrAspTrrMetProGluGlnProAsn 620
Db 2005 CTTGGAGTTAATGGTCTAATTTATAGATATATGATGATGATGATGATGATGATGATGAT 2064
Qy 621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
Db 2065 ATATTCCAGTGAGCAATTTGGAACCGCTTGAACGAGAAATTCACAGCTTTAAGACTGT 2124
Qy 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 2125 TTGTGTCCCACTGTTAGCGCTTTGTTTCCCTCATCTTTGTGTGGGAGTCTGATATCCAT 2184
Qy 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
```

Db 2185 GTGGATGCCAACGGCATTGATAACGTGGAGATGCT 2220

RESULT 9

ABV23741
ID ABV23741 standard; cDNA; 3498 BP.
XX
AC ABV23741;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 23732.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JB;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 4356-4357; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 3498 BP; 1103 A; 609 C; 736 G; 1040 T; 0 U; 10 Other;

Alignment Scores:

Pred. No.:	0	Length:	3498
Score:	3509.00	Matches:	671
Percent Similarity:	99.85%	Conservative:	0
Best Local Similarity:	99.85%	Mismatches:	1
Query Match:	99.21%	Indels:	1
DB:	5	Gaps:	0

US-10-047-855-3 (1-672) x ABV23741 (1-3498)

Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20

Db 221 ATGACACCTTCTCAGGTGCTTTGAAATTAAGAGGAACTCTTTTACCAGAGAGATTTT 280

Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrrAsnProGlnAsnAlaValAlaLeu 40

Db 281 GCGATATGTGGAAGCTGTGATGCTTTGGGAACTGGAAATCCTCAAAATGCTGTGGCTCTT 340
Qy 41 LeuProGluAenAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 341 CTTCCAGAGATGACACAGGGTGAAGCATGTATGGAAAGCAACCATTTGTACTCAGTAGA 400
Qy 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 401 GGAGTATCAGTTTCAGTATCGCTACTTCAAGGGGTACTTTTTAGAACCAAGACTATCGGT 460
Qy 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 461 GGTCCATGTCGAAGTGTAGTTTCAAGTGGGAGACTCATCTPACCAACCAAGTCAATTAACC 520
Qy 101 ProLeuGluSerGluIleIleAaspAspGlyGlnPheGlyIleHisAasnGlyValGlu 120
Db 521 CCTTTAGAAAGCAATTAATTATTGACGATGGACAAATTTTGGAAATCCCAATGGTGTGA 580
Qy 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
Db 581 ACTCTGGATTTCTGGATGCTGACATGTGACACTGAAATTAAGATTACGTTTGCATATTCT 640
Qy 141 GlutysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
Db 641 GAAAAACCTCTGTGTCAATTAACCAAGAAATAATTAATAAATCTAGATTTAGGGTGAAG 700
Qy 161 LeuThrLeuGluGlyLeuGluAspAspArgValSerProThrValLeuHis 180
Db 701 CTGACACTAGAAAGGCTGGAGAGATGACGATGATAGGGTATCTCCCACTGTACTCCAC 760
Qy 181 LysMetSerAenSerLeuGluLeuSerLeuLeuSerAenGluPheLysCysArgHis 200
Db 761 AAAATGTCCAATAGCTTGGAGATATCCTTAAGCGCAATAGTTCAGTTCAGGTGCGGCAT 820
Qy 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
Db 821 TCACAGCGGAGTGTGGTATGGCTTGCAGCCGTGATCGTGGACAGAGTACAGCATACAG 880
Qy 221 ThrMetGluProAspAasnLeuGluLeuPheAspPhePheGluGluAspLeuSerGlu 240
Db 881 AGATGGAAACCAAGATAACCTGGAATCTTCTTGTATTTTCGAAGAAGATCTCAGTGAG 940
Qy 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db 941 CACGTAGTTACGGTGATGCCCTTCTGGACATGTGGGTACAGCTTGCTCTTATCATCC 1000
Qy 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAen 280
Db 1001 ACCATTGCTGAGAGTGGAAAGAGTGTGGAATTTCTTACTCTTCCCATCATGAGCGAAAT 1060
Qy 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300
Db 1061 TCCCGGAAACCAATAGGCAAGTGGAGTTGACTATATTAATTTAAGCCATTACCGGA 1120
Qy 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
Db 1121 TACAGTTGTGACATGAAATCTTCAATTTCCAAAGTATTGGAAGCAAGAAATACCATGGAT 1180
Qy 321 ValGlyHisArgGlyAlaGlyAenSerThrThrThrAlaGlnLeuAlaLysValGlnGlu 340
Db 1181 GTTGGCCATCGAGTGCAGGAAACTCTACAACAACCTGCCCGAGCTGGCTTAAGTTGAC 1240
Qy 341 AenThrIleAlaSerLeuArgAenAlaAaSerHisGlyValAlaPheValGluPheAsp 360
Db 1241 AATACTATTGCTTCTTTAAGAAATGCTGTAGTCATGGGTGCAGCCCTTTGTAGAAATTGAC 1300
Qy 361 ValHisLeuSerLysAspPheValProValTyrHisAspLeuThrCysCysLeuThr 380
Db 1301 GTACACTTTTCAAGGACTTTGTGCCGTGGTGTATATCATGATCTTACCTGTTGTTGACT 1360
Qy 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400

Db 1361 ATGAA-AAGAAATTTGATGCTGATCCAGTTGAAATTTATTTGAAATTTCCAGTAAAGAAATTA 1419
Qy 401 ThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
Db 1420 ACAATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCCTGAAATCTAAGGAT 1479
Qy 421 ArgLysGluSerValValGlnGluGluAenSerPheSerGluAenGlnProPheProSer 440
Db 1480 CGGAAAGAAATCTGTGTTTCAAGGAGAAATTCCTTTTTCAGAAATCAGCCATTTCTCTCT 1539
Qy 441 LeuLysMetValLeuGluSerLeuProGluLysValGlyPheAenIleGluIleLysTrp 460
Db 1540 CTTAAGATGTTTTAGAGTCTTTTCCAGAGAGATGTAGGGTTTAAACATTTGAAATTAATAAG 1599
Qy 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAasnLeuSerThrTyrPheAspMetAasn 480
Db 1600 ATCTGCCAGCAAGAGGATGGAATGTGGATGGTAACCTTATCAACATATTTTGACATGAAT 1659
Qy 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAenSerGlyLysArgArgIleVal 500
Db 1660 CTGTTTTTGCATATAATTTTAAAAACTGTTTTAGAAAAATTTCTGGGAAAGAGAGAAATAGTG 1719
Qy 501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAenLysTyrPro 520
Db 1720 TTTTCTTCAATTTGATGTCAGATATTTGCACAAATGGTTCCGCAAAAGAGCAAAATATCCG 1779
Qy 521 IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer 540
Db 1780 ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTGAACTCATGGACCTCAGATCT 1839
Qy 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAenLeuLeuGlyIleAasnVal 560
Db 1840 CGGACAAACCCCATTTGCAATGAGCTTTGCACAGTTTGAAATCTTACTGGGATAAATGTA 1899
Qy 561 HisThrGluAspLeuLeuArgAasnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
Db 1900 CATACTGAAGACTTTGCTCAGAAACCCATCTCTATATTTCAAGAGGCAAAAGCTTAAGGACTA 1959
Qy 581 ValIlePheCysTrpGlyAspAspThrAasnAspProGluAenArgArgLysLeuLysGlu 600
Db 1960 GTCATATTTCTGCTGGGTGATGATACCAATGATCTGAAAAACAGAGAAATTTGAAGAA 2019
Qy 601 LeuGlyValAenGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAen 620
Db 2020 CTTGAGATTAAATGCTCTAAATTTATGATAGGATATATGATGATGCTGAAACCAACAAAT 2079
Qy 621 IlePheGlnValGluGlnLeuGluArgLysGlnGluLeuProGluLeuLysSerCys 640
Db 2080 ATATTCCAAGTGGAGCAATTTGGAACGCTGAAGCAGGAAATTTGCCAGAGCTTTAAGAGCTGT 2139
Qy 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 2140 TTGTGTCCCACTGTGTAGCCGCTTTGTCCCTCATCTTTGTGTGGGAGTCTGATATATCAT 2199
Qy 661 ValAspAlaAenGlyIleAspAasnValGluAasnAla 672
Db 2200 GTGGATGCCAAGCATTTGATAACCTGGAGAAATGCT 2235
RESULT 10
ADJ57938
ID ADJ57938 standard; cDNA; 3381 BP.
XX
AC ADJ57938;
XX
DT 06-MAY-2004 (first entry)
XX
DE Rat NARC 16 cDNA.
XX
KW Neuronal apoptosis regulated candidate; NARC; diagnosis; cancer;
KW lung disease; cirrhosis; hepatitis; atherosclerosis;
KW myocardial infarction; inflammation; anaemia; glomerulonephritis;
KW osteoporosis; AIDS; acquired immunodeficiency syndrome;
KW Parkinson's disease; Alzheimer's disease; stroke; dermatitis;

drug screening; gene therapy; cytostatic; hepatotropic; neurotropic; cerebrotropic; dermatological; virucide; neuroprotective; phosphatidylglycerol phosphate synthase; PGP synthase; rat; ss.

Rattus norvegicus.

US2004009553-A1.

15-JAN-2004.

30-APR-2003; 2003US-00426776.

27-SEP-1999; 99US-00405045.

22-OCT-1999; 99US-0161188P.

31-JAN-2000; 2000US-00495823.

28-FEB-2000; 2000US-0185517P.

20-OCT-2000; 2000US-00692785.

31-JAN-2001; 2001US-00773426.

28-FEB-2001; 2001US-00795691.

31-OCT-2001; 2001US-0335003P.

31-OCT-2001; 2001US-0335037P.

25-MAR-2002; 2002US-00105992.

28-AUG-2002; 2002US-00229662.

30-OCT-2002; 2002US-00284014.

30-OCT-2002; 2002US-00284059.

09-DEC-2002; 2002US-00314881.

(MILL-) MILLENNIUM PHARM INC.

Glucksmann MA, Williamson MJ, Tsai P, Rudolph-Owen LA;

Kapeller-Libermann R, Meyers RE, Chiang LW, Hunter JJ;

WPI; 2004-090469/09.

New isolated nucleic acid molecules and polypeptides (e.g. 27411, 23413, 22436, 23553, NARC SCI or NARC I) useful for diagnosing, preventing or treating disorders associated with the protein, e.g. cancer, atherosclerosis or AIDS.

Claim 1; SEQ ID NO 39; 260pp; English.

The present invention provides isolated nucleic acid molecules and proteins designated 27411, 23413, 22438, 23553, 25278, 26212, NARC SCI, NARC 10A, NARC 1, NARC 12, NARC 13, NARC 17, NARC 25, NARC 3, NARC 4, NARC 7, NARC 8, NARC 11, NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20, NARC 26, NARC 27, NARC 28, NARC 30, NARC 5, NARC 6, NARC 9, NARC 10C, NARC 8B, NARC 9, NARC 2A, NARC 16B, NARC 1C, NARC 1A, NARC 25, 86604 or 32222. The invention is useful in diagnosing, preventing or treating disorders such as cancer, lung diseases, cirrhosis, hepatitis, atherosclerosis, myocardial infarction, inflammation, anaemia, glomerulonephritis, osteoporosis, AIDS (acquired immunodeficiency syndrome), Parkinson's disease, Alzheimer's disease, stroke and dermatitis. These may also be used in drug screening. The invention is also useful in gene therapy. The present sequence is rat neuronal apoptosis regulated candidate (NARC) cDNA.

Sequence 3381 BP; 990 A; 689 C; 729 G; 973 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.58e-291	Length:	3381
Score:	2965.50	Matches:	571
Percent Similarity:	90.76%	Conservative:	38
Best Local Similarity:	85.10%	Mismatches:	33
Query Match:	83.84%	Indels:	29
DB:	12	Gaps:	6

US-10-047-855-3 (1-672) x ADJ57938 (1-3381)

QY	1	Met-Thr-Pro-Ser-Gln-Val-Ala-Phe-Glu-Ile-Arg-Gly-Thr-Leu-Leu-Pro-Gly-Glu-Val-Phe	20
DB	186	ATGACACCTTCTCAGGTACCTTTGAATAAAGGAACCTCTTTTACGAGAGGCTTT	245
QY	21	Ala-Ile-Cys-Gly-Ser-Cys-Asp-Ala-Leu-Gly-Asn-Trp-Asn-Pro-Gln-Ala-Val-Ala-Leu	40

DB	246	GCAATGTTGGAAACCTGATCCTTGGAAATCGAGTCTCAAATGCTGTGCTCT	302
QY	41	Leu-Pro-Glu-Asn-Asp-Thr-Gly-Glu-Ser-Met-Leu-Tyr-Lys-Ala-Thr-Ile-Val-Leu-Ser-Arg	60
DB	303	CTTACTGAGAGTGAGCAGCGGAAAGT---GTATGGAAAGCAGTGATGTTCTTAGTAGA	359
QY	61	Gly-Val-Ser-Val-Gln-Tyr-Arg-Tyr-Phe-Lys-Gly-Tyr-Phe-Leu-Glu-Pro-Lys-Thr-Ile-Gly	80
DB	360	GGAAATGTCGTTGAAGTACCGCTACTTCAGAGGCTGCTTTTATAGAACCAAGACTATCGGT	419
QY	81	Gly-Pro-Cys-Gln-Val-Ile-Val-His-Tyr-Trp-Glu-Thr-His-Leu-Gln-Pro-Arg-Ser-Ile-Thr	100
DB	420	GGTCCATGTCAAGTCATAGTTCACAAGTGGGAGACTCATCTAACACGACGATCAATAACC	479
QY	101	Pro-Leu-Glu-Ser-Glu-Ile-Ile-Leu-Asp-Gly-Gln-Phe-Gly-Ile-His-Asn-Gly-Val-Glu	120
DB	480	CTTTAGAAAACGAATCATTTATGACGATGACAAATTTGGAAATCCACATGGTGTGAA	539
QY	121	Thr-Leu-Asp-Ser-Gly-Trp-Leu-Thr-Cys-Gln-Thr-Glu-Ile-Arg-Leu-Arg-Leu-His-Tyr-Ser	140
DB	540	ACACTGGATTCTGGATGGCTTACCTGTGACAGTGAATAAGACTGCGTCTGCAATTTTCT	599
QY	141	Glu-Lys-Pro-Pro-Val-Ser-Ile-Thr-Lys-Lys-Leu-Lys-Lys-Ser-Arg-Phe-Arg-Val-Lys	160
DB	600	GAGAAACCTCTCTGTTTCAATTACCAAGAAAAAGTTCAAAAAATCTAGATTTAGGGTAAAG	659
QY	161	Leu-Thr-Leu-Glu-Gly-Leu-Glu-Glu-----Asp-Asp-Asp-Asp-Arg-Val-Ser-Pro-Thr-Val	178
DB	660	CTTACACTAGAGGGTCTGGAGGAAGATGATGACGAGATGATAAGGACATCTCCACTGTT	719
QY	179	Leu-His-Lys-Met-Ser-Asn-Ser-Leu-Glu-Ile-Ser-Leu-Ile-Ser-Asp-Asn-Glu-Phe-Lys-Cys	198
DB	720	CTTCAACAGATGCTCCAAATAGCTGGAGATATCTTAAATAGTGACAAATGAGTTCAGTGC	779
QY	199	Arg-His-Ser-Gln-Pro-Glu-Cys-Gly-Tyr-Gly-Leu-Gln-Pro-Asp-Arg-Trp-Thr-Glu-Tyr-Ser	218
DB	780	AGGCACCTCACAGCCAGATGTGGTATGGCTTTACAGCCCTGACCGCTGGACAGAGTACAGC	839
QY	219	Ile-Gln-Thr-Met-Glu-Pro-Asp-Asn-Leu-Glu-Ile-Phe-Asp-Phe-Phe-Glu-Glu-Asp-Leu	238
DB	840	ATACAGACAAATGGAGCGGACAACTTGAACCTCACTTTTGACTTTTTTGGAGGAAGATCTC	899
QY	239	Ser-Glu-His-Val-Val-Gln-Gly-Asp-Ala-Leu-Pro-Gly-His-Val-Gly-Thr-Ala-Cys-Leu-Leu	258
DB	900	AGTGAGCATGTAGTCCAGGGTGTGTTCTTCTCGGACATGTGGGACACAGCATGCGCTCCTG	959
QY	259	Ser-Ser-Thr-Ile-Ala-Glu-Ser-Gly-Lys-Ser-Ala-Gly-Ile-Leu-Thr-Leu-Pro-Ile-Met-Ser	278
DB	960	TCATCTCATTGCTGAGAGTGAGAGAGCGCTGGAATCCTTACTCTTCCCATTCAATGAGC	1019
QY	279	Arg-Asn-Ser-Arg-Lys-Thr-Ile-Gly-Lys-Val-Arg-Val-Asp-Tyr-Ile-Ile-Ile-Lys-Pro-Leu	298
DB	1020	AGAAGTCCAGAAAACTATAGCGAAAGTCAGAGTTGATTTTATCATCATCAAGCCATT	1079
QY	299	Pro-Gly-Tyr-Ser-Cys-Asp-Met-Lys-Ser-Ser-Phe-Ser-Lys-Tyr-Trp-Lys-Pro-Arg-Ile-Pro	318
DB	1080	CCAGGATATAGTTGTTCTATGTCAGTCTTCACTTCTCCAAGTATTTGGAACCAAGAAATACCA	1139
QY	319	Leu-Asp-Val-Gly-His-Arg-Gly-Ala-Gly-Asn-Ser-Thr-Thr-Thr-Ala-Gln-Leu-Ala-Lys-Val	338
DB	1140	CTGGATTTGGACATCGTGGTGCAGGGAACCTCAACAACTGCCAACGCTGGCTAAAGTA	1199
QY	339	Gln-Glu-Asn-Thr-Ile-Ala-Ser-Leu-Arg-Asn-Ala-Ser-His-Gly-Ile-Ala-Phe-Val-Glu	358
DB	1200	CAGGAAAAATACATTGCTTCTTTAAGAAATGTCGACGACCATGGTGGTGGTGGAA	1259
QY	359	Arg-Asp-Val-His-Leu-Ser-Lys-Asp-Phe-Val-Pro-Val-Val-Tyr-His-Asp-Leu-Thr-Cys-Cys	378
DB	1260	TTTGATGTCACCTTTCAAGGACTTAGTGCTCTAGTGTATCATGATCTCACCTGCTGT	1319
QY	379	Leu-Thr-Met-Lys-Lys-Phe-Asp-Ala-Asp-Pro-Val-Glu-Leu-Phe-Glu-Ile-Pro-Val-Lys	398

Db 1320 TTAACATATGAAAGAAATATGAGCTGATCCAGTTGAATGTTTGAATCCAGTAAAG 1379
 Qy 399 GluLeuThrPheAspGlnLeuGlnLeuLeuLeuLeuLeuHisValThrAlaLeuLysSer 418
 Db 1380 GAATTAACATTCGACCAACTCCAGTTATTGAGCTTTCTCATGTGACTGCACATAAAACC 1439
 Qy 419 LysAspArgLysGluSerValValGlnGlnGlnSerPheSerGluAenGlnProphe 438
 Db 1440 AAAGACCAGAAACAATGTATGGCTGAGGAGGAAATTCCTTTCTGAAACCAACCATTT 1499
 Qy 439 ProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIle 458
 Db 1500 CCTTCTCTTAAGATGGTTTAGAGTCATTCGCCAGAAATATGAGGATTTAATATAGAATA 1559
 Qy 459 LysTrpIleCysGlnGlnAtqAspGlyMetTrpAspGlyAsnLeuSerThrTyPheAsp 478
 Db 1560 AAATGGATTTCCCAACACAGGAGTGGAGTATGGGACGGCAACTTATCGACATATTTTGTAT 1619
 Qy 479 MetAsnLeuPheLeuAspIleLeuLysThrValLeuGluAenSerGlyLysArgArg 498
 Db 1620 ATGAATGCATTTTGGATATAATTTTAAAAACTGTTTAGAAAAATTCGGGAGAGGAGA 1679
 Qy 499 IleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLys 518
 Db 1680 ATAGTATTTTCTTCATTTGTCGACATCTGTACATGGTTCCGACAGAAACAAACAAA 1739
 Qy 519 TyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyProGluLeuMetAspLeu 538
 Db 1740 TATCCCATATTTTTCACCAAGGAAGCTGCATTTACCTTGAACCTCATGGACCTC 1799
 Qy 539 ArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIle 558
 Db 1800 AGATCTCGGACACACCCATTGCAATGAGCTTTGCACAGTTTGAATAATTTTGGGGATA 1859
 Qy 559 AsnValHisThrGluAspLeuLeuArgAsnProSerTyIleGlnGluAlaLysAlaLys 578
 Db 1860 AATGCCCATCTGAAGATCTCTCTAGAAACCCATCTATGTCCAAAGAGGCAAAAGATAG 1919
 Qy 579 GlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgLysLeu 598
 Db 1920 GGATTTGGTATATTTCTGCTGGGGTGATGTATACATGATCTCTGAAAACAGAGGAAACTG 1979
 Qy 599 LysGluLeuGlyValAsnGlyLeuLeuTyAspArgGlyLeuLeuLeuLeuLeuLeuGln 618
 Db 1980 AAGGAATTTGGATTAATGCTCTAATATATGATAGTAT----- 2018
 Qy 619 ProAsnIlePheClnValGlnLeuGluArgLeuLysGlnGluLeu----- 634
 Db 2019 -----TTGTTTTTTGTAATAAATCTCCATGGAAATTTGTTCAACACAGTGTAGTTTATCTA 2072
 Qy 635 -----ProGluLeuLysSer-----CysLeuCys 642
 Db 2073 TTTTAACTATTTTAAATTTAGATAGTTTGGCTTAAAGTTTATCTTGACACTGTGACCTT 2132
 Qy 643 ProThrValSerArgPheValProSerSerLeu 653
 Db 2133 TCCAGGTGTGAGATATGTCAAAAGCCACTTA 2165
 RESULT 11
 ABA09667
 ID ABA09667 standard; DNA; 2477 BP.
 XX
 AC ABA09667;
 XX
 DT 15-JAN-2002 (first entry)
 XX
 DE Human bone marrow expressed oligonucleotide SEQ ID NO: 176.
 KW Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnery;
 KW antinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
 KW antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
 KW antitumor; fungicide; antidiabetic; antiaesthetic; antiallergic;
 KW immunostimulant; analgesic; cerebroprotective; antianaemic; infection;

KW nervous system disorder; autoimmune disorder; inflammation; allergy; ds.
 XX Homo sapiens.
 OS WO200174836-A1.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US010472.
 PF 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 PP 30-NOV-2000; 2000US-0250583P.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Ford JE, Boyle BJ;
 PI WPI; 2001-626375/72.
 DR New bone marrow-expressed nucleic acids and polypeptides, useful for
 XX diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
 PT and increasing hematopoiesis, stem cell survival and bone growth and
 PT remodeling.
 XX Claim 1; Page 254-255; 380pp; English.
 PS The present invention relates to bone marrow expressed polynucleotides
 XX and proteins. These sequences can be used in the treatment of
 CC inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
 CC and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
 CC disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
 CC cell disorders, platelet disorders, stem cell disorders, bone
 CC degenerative disorders, autoimmune disorders, for example multiple
 CC sclerosis, diabetes and arthritis, viral and bacterial infections,
 CC allergies and blood coagulation disorders. The present sequence is a DNA
 CC of the invention
 XX
 SQ Sequence 2477 BP; 681 A; 500 C; 642 G; 654 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7.42e-278 Length: 2477
 Score: 2829.50 Matches: 573
 Percent Similarity: 79.97% Conservative: 14
 Best Local Similarity: 78.07% Mismatches: 36
 Query Match: 80.00% Indels: 111
 DB: 5 Gaps: 11
 US-10-047-855-3 (1-672) x ABA09667 (1-2477)
 Qy 18 GluValPheAlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAla 37
 Db 369 GAAATTTTGGCATATGTGGAAGCTGTATCTTTGGGAAACTGGAATCTCTCAAAATGCT 428
 Qy 38 ValAlaLeuLeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleVal 57
 Db 429 GTGGCTCTTCTCCAGAGAATGACACAGTGAAGCATGCTATGGAAGCAACCATTTGTA 488
 Qy 58 LeuSerArgGlyValSerValGlnTrpArgTyPheLysGlyTyPheLeuGluProLys 77
 Db 489 CTCAGTAGAGGATGATCAGTTTCAGTATCGTACTCTCAAGGGTACTTTTTAGAACCAAG 548
 Qy 78 ThrIleGlyGlyProCysGlnValIleValHisLysTrpGluThrHisLeuGln---Pro 96
 Db 549 -----GAAACATTCATCATCGGTGCTTCTTGTGTAACTTTCCA 590
 Qy 97 ArgSerIleThrProLeuGluSerGluIleIleAspAspGlyGlnPhe----- 113
 Db 591 AGTTCT-----TCAGTTCTCTTGTTCAGACTGGACAGTTCTCTGGGAGG 638
 Qy 113 ----- 113

Db 639 GACATAGACAGACCAAACTAAGTCTCTGGCGAGGGCTGGGGTTTGAATGGGCT 698
Qy 114 -----Gly11eHisAsnGly 118
Db 699 ATAGAGCTATGCATGGCAGTCAAAGAGGATGTTTCGACAGGAAGTAGGACCATATAGG 758
Qy 119 ---valGluThrLeuAspSerGlyTrpLeuThrCysGln-----ThrGlu 132
Db 759 TTACTTCTCGATGTTCCCATGGCAATTTGTAACACTGTCGTGGCACTGATGGAGTGTAGCA 818
Qy 133 IleArgLeu---ArgLeuHis----- 138
Db 819 GTGAGATGACAGAGGTCACTCTCATTCGCCATCTTGGTTTGCCTACTGTGCTAGTGGG 878
Qy 139 TyrSerGluLysProValSer----- 146
Db 879 TTCTCCTTGGAGCCTGTGTGTAATAATGACTGTGGGGCCTCTTCTCGACAGGTGCAACAG 938
Qy 147 -----IleThrLysLysLysLeuLysLysSer-----ArgPheArg 158
Db 939 GGCTTCGTGTTTCATCATCTGCCTCCTCTCTCTCATATTGCACTGAGGCCAAAGG 998
Qy 159 ValLysLeuThrLeuGluGlyLeuGluAspAspAspArgValSerProThrVal 178
Db 999 GTGAAGCTGACACTAGAAGCCTGGAGAGATGACGATGATAGGGTATCTCCCACTGTA 1058
Qy 179 LeuHisLysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCys 198
Db 1059 CTCACAAATGTCCAAATAGCTTGGAGATATCTTAATAAGGACCAATGAGTTCAGTGC 1118
Qy 199 ArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSer 218
Db 1119 AGGCATTTCAGCGCGAGTGTGTTATGCTTGCAGCCTCATCGTGGGACAGAGTACAGC 1178
Qy 219 IleGlnThrMetGluProAspAsnLeuGluIlePheAspPheGluGluAspLeu 238
Db 1179 ATACAGACCATGGAACCATACCTCGGAACATAATCTTTGATTTTTTCGAAGAAGATCTC 1238
Qy 239 SerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeu 258
Db 1239 AGTGACACGTAGTTCAGGGTGATGCTCTCTCGGACATGTCGGTACAGCTGTCTCTTA 1298
Qy 259 SerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSer 278
Db 1299 TCATCCACCATGCTCAGAGTGGAAAGAGTGTGGAAATCTTACTCTTCCCATCATGAGC 1358
Qy 279 ArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeu 298
Db 1359 AGAAATTTCCCGGAAAAACAATAGCAAGTGAAGTGTGACTATATAATTTAAGCCATTA 1418
Qy 299 ProGlyTyrSerCysAspMetLysSerPheSerLysTyrTrpLysProArgIlePro 318
Db 1419 CCAGGATACAGTTGTGACATGAATCTTCATTTCCAGTATTTGGAGCCAGNATACCA 1478
Qy 319 LeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysVal 338
Db 1479 TTGGATGTTGGCCATCGAGTGCAGGAAACTCTACAAACACTGCCCCAGCTGGCTAAAGTT 1538
Qy 339 GlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGlu 358
Db 1539 CAAGAAATACATATTGCTTCTTTAAGAAATGCTGTAGTCATGGTGCAGCCTTTGAGAA 1598
Qy 359 PheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCys 378
Db 1599 TTTGAGTACACTTTCAAAGGACTTTGTGGCCGTGTATATCATGATCTTACCTGTGTT 1658
Qy 379 LeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLys 398
Db 1659 TTGACTATCAAAAGAAATTTGATGCTGATCCAGTTGAATTTTGAATTCAGTAAAA 1718
Qy 399 GluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSer 418
Db 1719 GAATTAACATTTGACCAACTCCAGTTGTTTAAAGCTCACTCATGTGACTGCACTGAATCT 1778

Qy 419 LysAspArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProphe 438
Db 1779 AAGGATCGGAAGAACTCTGTGTTTCAGGAGGAAATTCCTTTTCAGAAAATCAGCCATTT 1838
Qy 439 ProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIle 458
Db 1839 CCTTCTTTAAGATG----- 1853
Qy 459 LysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTrpPheAsp 478
Db 1854 -----CATGGAATGTGGATGGTAACTTATCAACATATTTTGAC 1892
Qy 479 MetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArg 498
Db 1893 ATGAATCTGTTTTCGATATAATTTTAAAAACTGTTTTAGAAAAATCTCGGAAGAGCAGA 1952
Qy 499 IleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLys 518
Db 1953 ATAGTGTGTTTCTTCATTTGATGCAGATATTTGCACAAATGTTTCGGCAAAAGCAGAA 2012
Qy 519 TyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeu 538
Db 2013 TATCGATACATATTTTAACTCAAGGAAATCTGAGATTTATCTGAACTCATGGACCTC 2072
Qy 539 ArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIle 558
Db 2073 AGATCTCGGACAAACCCCATTCGAATGAGCTTTGCACAGTTCGMAAATCTACTGGGGATA 2132
Qy 559 AsnValHisThrGluAspLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLys 578
Db 2133 AATGTACATACCTGAAGACTTGTCTAGAAACCCCATCTATATTTCAAGAGGCAAAAGCTAAG 2192
Qy 579 GlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeu 598
Db 2193 GGACTAGTCATATTTCTGTGGGGTGATGATACCAATGATCTCTGAAAAACAGAGAAATTTG 2252
Qy 599 LysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGln 618
Db 2253 AAGGAACTTGGAGTTAATGGTCTAATTTATGATAGATATATGATTTGGATGCTGAACAA 2312
Qy 619 ProAsnIlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLys 638
Db 2313 CCAAAATATATTCGAAGTGGAGCAATTTGGAACCCCTGAAGCAGAGAAATTCAGAGCTTAAAG 2372
Qy 639 SerCysLeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAsp 658
Db 2373 AGCTGTTGTGTGCCCATGTTAGCCGCTTTGTTCCCTCATCTTTGTGGGGAGTCTGAT 2432
Qy 659 IleHisValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2433 ATCCATGTGGATGCCAACGGCATTCGTAACGTGGAGATGCT 2474

RESULT 12

AAD06007

ID AAD06007 standard; DNA; 2738 BP.

XX AAD06007;

XX AAD06007;

DT 31-JUL-2001 (first entry)

XX Human neuronal apoptosis regulated candidate (NARC) 9B DNA.

Human; neuronal apoptosis regulated candidate 9B; NARC 9B; cytostatic;
chromosome mapping; gene therapy; antisense therapy; lung disorder;
central nervous system disorder; apoptosis; spleen disorder; angina;
tuberculosis; Goodpasture's syndrome; liver disorder; jaundice;
infectious disorder; brain disorder; cerebral oedema; gonorrhea;
heart disorder; kidney disorder; glomerulonephritis; testes; virucide;
epididymis disorder; skeletal muscle disorder; pancreatic disorder;
diabetes; cytoprotectant; immunostimulant; tumour; antiinflammatory;
antimicrobial; neuroprotective; gynaecological; ds.

OS Homo sapiens.
 XX WO200131007-A2.
 XX
 XX 03-MAY-2001.
 XX
 XX 20-OCT-2000; 2000WO-US029132.
 XX
 XX 22-OCT-1999; 99US-0161188P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Chiang LW;
 XX
 XX WPI; 2001-308641/32.
 DR
 XX
 XX Rat brain polypeptides, nucleic acids and antibodies, useful for
 PT diagnosis and treatment of central nervous system disorders and disorders
 PT associated with aberrant apoptosis.
 XX
 XX Claim 1; Page 146; 161pp; English.
 XX
 XX The invention relates to human homologues of neuronal apoptosis regulated
 CC candidate (NARC) nucleic acid molecules and proteins derived from rat
 CC brain and programmed cell death libraries. The nucleic acids of the
 CC brain and programmed cell death libraries. The nucleic acids of the
 CC invention are useful for assaying the presence of a nucleic acid molecule
 CC and for chromosome mapping. They are also used in gene therapy and
 CC antisense therapy. The NARC sequences are useful for treating central
 CC nervous system disorders and disorders involving aberrant apoptosis, for
 CC inducing an immune response and for isolating binding partners. Diseases
 CC treated include spleen disorders (e.g. tuberculosis and congestive
 CC splenomegaly), lung disorders (e.g. adult respiratory distress syndrome,
 CC Goodpasture's syndrome and bronchial asthma), liver disorders (e.g.
 CC jaundice and hepatic failure), infectious disorders (e.g. viral
 CC hepatitis), brain disorders (e.g. cerebral oedema, hypertensive
 CC encephalopathy and hydrocephalus), heart disorders (e.g. heart failure,
 CC angina and myocardial infarction), kidney disorders (e.g. cysts and
 CC glomerulonephritis), testes and epididymis disorders (e.g. gonorrhoea and
 CC syphilis), skeletal muscle disorders (e.g. tumours) and pancreatic
 CC disorders (e.g. pancreatitis and diabetes). The present sequence is human
 CC neuronal apoptosis regulated candidate (NARC) 9B DNA
 XX
 XX Sequence 2738 BP; 880 A; 450 C; 540 G; 868 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 1.76e-249 Length: 2738
 Score: 2552.00 Matches: 486
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 72.15% Indels: 0
 DB: 4 Gaps: 0

US-10-047-855-3 (1-672) x AAD06007 (1-2738)

Qy 187 GluIleSerLeuIleSerAspAenGluPheLysCysAerHisSerGlnProGluCysGly 206
 Db 18 GAGATATCTTATATAGCGCATAGTTCAGTGCAGGCAATTCACAGCGGAGTGTGT 77
 Qy 207 TyrGlyLeuGlnProAspArgfrrpThrGluTyrSerIleGlnThrMetGluProAspAen 226
 Db 78 TATGGCTTCAGCGCTGATCGTTGGACAGAGTACAGCATACAGACGATGGAAACAGATAAC 137
 Qy 227 LeuGluLeuIlePheAppPheGluGluAspLeuSerGluHisValGlnGlyAsp 246
 Db 138 CTGGAACATACTTGTGATTTTTCGAAGAAGATCTCAGTGAGGACGCTAGTTCAGGGTGTAT 197
 Qy 247 AlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAlaGluSerGly 266
 Db 198 GCCCTTCTCGACATGTGGGTACAGCTTGCTCTTATCATCCACCATTCGTGAGAGTGA 257
 Qy 267 LysSerAlaGlyIleLeuThrLeuProIleMetSerArgAenSerArgLysThrIleGly 286
 Db 258 AAGAGTGTGGAATTTCTTACTCTTCCCATCATGACGAGNAATTCCTCCGGAACATAGGC 317

287 LysValArgValAspTyrIleIleLysProLeuProGlyTyrSerCysAspMetLys 306
 318 AAAAGTGAGAGTTGACTATATATTAATTAAGCCATTACCCAGGATACAGTTGTGACATGAA 377
 307 SerSerPheSerLysTyrTrpLysProArgLleProLeuAspValGlyHisArgGlyAla 326
 378 TCTTTCATTTTCCAGTATTTGGAAGCCAAAGATACCATTTGGATGTTGGCCATCGAGGTGA 437
 327 GlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeu 346
 438 GGAACCTCTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 497
 347 ArgAsnAlaAlaSerHisGlyValAlaPheValGluPheAspValHisLeuSerLysAsp 366
 498 AGAAATGCTGCTAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
 367 PheValProValValThrHisAspLeuThrCysCysLeuThrMetLysLysLysPheAsp 386
 558 TTTGTGCCCGTGGTATATCATGATCTTACCTGTTGTTGCTATGATAAAGAAATTTGAT 617
 387 AlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAspGlnLeuGln 406
 618 GCTGATCCAGTTGAATTTTGAATTTCCAGTAAAGAAATTAACATTTGACCACTCCAG 677
 407 LeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSerValVal 426
 678 TTGTTAAAGCTCACTCATGTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
 427 GlnGluGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMetValLeuGlu 446
 738 CAGAGGAAATTTCTTTTTCAGAAATCAGCCATTTCTTCTTAAGATGTTTGTAGAG 797
 447 SerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGlnGlnArgAsp 466
 798 TCTTTGCCAGAGATGTAGGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 857
 467 GlyMetTrpAspGlyLeuSerThrTyrPheAspMetAenLeuPheLeuAspIleIle 486
 858 GGAATGTGGATGTGATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 917
 487 LeuLysThrValLeuGluAsnSerGlyLysArgGlyValPheSerSerPheAspAla 506
 918 TTAATAAATCTGTTTAAAGAAATTTCTGGAGAGAGAGATAGTGTCTTCTTCTTATTTGATGCA 977
 507 AspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeuThrGln 526
 978 GATATTTGCACATATGTTTGGGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1037
 527 GlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThrProIleAla 546
 1038 GGAATAATCTGAGATTTTATCTGAACTCATGACCTCAGATCTCGACACACCCCATTTGCA 1097
 547 MetSerPheAlaGlnPheGluAsnLeuGlyIleAsnValHisThrGluAspLeuLeu 566
 1098 ATGAGCTTTGCACAGTTTGAATAATCTACTGGGATATAATGTATACATCTGGAAGACTTGGCTC 1157
 567 ArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpGly 586
 1158 AGAAACCCATCTCTATTTTCAAGAGGCAAAAGCTAAGGAGACTAGTCTATTTCTGCTGGGGT 1217
 587 AspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyValAsnGlyLeu 606
 1218 GATGATACCAATGATCTCTGAAACAGAGAGAAATTTGAAGGAACCTTGGAGTTTAATGCTCTA 1277
 607 IleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePheGlnValGluGln 626
 1278 ATTTATGATAGGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1337
 627 LeuGluArgLeuLysGlnLeuProGluLeuLysSerCysLeuCysProThrValSer 646
 1338 TTGGAAACGCTTGAAGCAGGAATTTGCCAGAGCTTAAAGAGCTGTTTGTGTGCTCCACTGTAGC 1397

QY 647 ArgPheValProSerSerLeuGlyGlySerAspIleHisValAspAlaAsnGlyIle 666
 Db 1398 CGTTTTGTTCCCTCATCTTTGTGTGGGAGTCTGATATCCATGCGATGCCAAGCGCAT 1457
 QY 667 AspAsnValGluAsnAla 672
 Db 1458 GATACGTTGGAGATGCT 1475

RESULT 13

ADJ57931
 ID ADJ57931 standard; cDNA; 2738 BP.

AC ADJ57931;

DT 06-MAY-2004 (first entry)

XX Human NARC 9 cDNA.

XX Neuronal apoptosis regulated candidate; NARC; diagnosis; cancer;
 KW lung disease; cirrhosis; hepatitis; atherosclerosis;
 KW myocardial infarction; inflammation; anaemia; glomerulonephritis;
 KW osteoporosis; AIDS; acquired immunodeficiency syndrome;
 KW Parkinson's disease; Alzheimer's disease; stroke; dermatitis;
 KW drug screening; gene therapy; cytostatic; hepatotropic; neurotropic;
 KW cerebroprotective; dermatological; virucide; neuroprotective;
 KW phosphatidyglycerolphosphate synthase; PGP synthase; human; ss.

XX Homo sapiens.

OS US2004009553-A1.

PN 15-JAN-2004.

PD 30-APR-2003; 2003US-00426776.

PF 27-SEP-1999; 99US-00406045.

PR 22-OCT-1999; 99US-0161188P.

PR 31-JAN-2000; 2000US-00495823.

PR 28-FEB-2000; 2000US-0185517P.

PR 20-OCT-2000; 2000US-00692785.

PR 31-JAN-2001; 2001US-00773426.

PR 28-FEB-2001; 2001US-00795691.

PR 31-OCT-2001; 2001US-0335003P.

PR 31-OCT-2001; 2001US-0335037P.

PR 25-MAR-2002; 2002US-00105992.

PR 28-AUG-2002; 2002US-00229662.

PR 30-OCT-2002; 2002US-00284014.

PR 30-OCT-2002; 2002US-00284059.

PR 09-DEC-2002; 2002US-003114881.

XX (MILL-) MILLENNIUM PHARM INC.

XX Glucksman MA, Williamson MJ, Tsai F, Rudolph-Owen LA;

PI Kapeller-Libermann R, Meyers RE, Chiang LW, Hunter JJ;

XX WPI; 2004-090469/09.

XX New isolated nucleic acid molecules and polypeptides (e.g. 27411, 23413,

PT 22438, 23553, NARC SCI or NARC 1) useful for diagnosing, preventing or

PT treating disorders associated with the protein, e.g. cancer,

PT atherosclerosis or AIDS.

XX Claim 1; SEQ ID NO 32; 260pp; English.

CC glomerulonephritis, osteoporosis, AIDS (acquired immunodeficiency
 CC syndrome), Parkinson's disease, Alzheimer's disease, stroke and
 CC dermatitis. These may also be used in drug screening. The invention is
 CC also useful in gene therapy. The present sequence is human neuronal
 CC apoptosis regulated candidate (NARC) cDNA.

XX Sequence 2738 BP; 880 A; 450 C; 540 G; 868 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.76e-249 Length: 2738
 Score: 2552.00 Matches: 486
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 72.15% Indels: 0
 DB: 12 Gaps: 0

US-10-047-855-3 (1-672) x ADJ57931 (1-2738)

QY 187 GluIleSerLeuIleSerAspAsnGluPheLysCysArgHisSerGlnProGluCysGly 206
 Db 18 GAGATATCCTTAATAAGCCGACCAATGAGTTCAAGTGCAGGCATTCACAGCCGGAGTGTGGT 77
 QY 207 TyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGlnThrMetGluProAspAsn 226
 Db 78 TATGGCTTCAGCCTGATCGTTGGACAGAGTACAGCATACAGACATGGAACAGATAC 137
 QY 227 LeuGluLeuIlePheAspPhePheGluGluAspLeuSerGluHisValGlnGlyAsp 246
 Db 138 CTGGAACATACTTTGATTTTTTCGAAGAAGATCTCAGTGAGCACCTAGTTTCAGGGTGAT 197
 QY 247 AlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAlaGluSerGly 266
 Db 198 GCCCTTCTCGACATGTGGGTACAGCTTGTCTTATCATCCACCATTCAGAGGTGGA 257
 QY 267 LysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsnSerArgIleGly 286
 Db 258 AAGATGCTGGAAATCTTACTCTTCCATCATGAGCAGAAATTCGCGGAAAAAATAGGC 317
 QY 287 LysValArgValAspTyrIleIleLysProLeuProGlyTyrSerCysAspMetLys 306
 Db 318 AAGTGAGAGTTGACATATATATTAAAGCCATTACAGGATACAGTTGTGACATGAAA 377
 QY 307 SerSerPheSerLysTyrTrpLysProArgIleProLeuAspValGlyHisArgGlyAla 326
 Db 378 TCTTCATTTTCCAAATATGGAAGCCAGAAATACCATTCGATGTCGCCATCAGGTGCA 437
 QY 327 GlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeu 346
 Db 438 GGAAACTCTACAACTGCCAGCTGGCTAAAGTTCAAGAAATATCTATTGCTTCTTTA 497
 QY 347 ArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAsp 366
 Db 498 AGAATGCTGCTAGTATGTCAGCTTTGTAGAAATTTGACGTACACCTTTCAAGGAC 557
 QY 367 PheValProValValTyrHisAspLeuThrCysCysLeuThrMetLysLysLysPheAsp 386
 Db 558 TTGTGCGCGTGATATCATGATCTTACCTGTGTGTGCTATGAAAGAAATTTGAT 617
 QY 387 AlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAspGlnLeuGln 406
 Db 618 GCTGATCCAGTTGAATTTTGAATTTCCAGTAAAGAAATTAACATTTGCCAACCTCCAG 677
 QY 407 LeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSerValVal 426
 Db 678 TTGTTAAAGCTCACTCATGTGCTGCACTGCAATCTAAGATCGGAAGAAATCTGTGTT 737
 QY 427 GlnGluGluAsnSerPheSerGluAsnGlnProPheProSerSerLeuLysMetValLeuGlu 446
 Db 738 CAGGAGGAAAAATCTCTTTTCAGAAAAATCAGCAATTCCTTCTTAAAGATGGTTTAGAG 797
 QY 447 SerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGlnGlnArgAsp 466
 Db 798 TCTTTGCCAGAGATGTAGGTGTTTAAATTTGAATTAATTTGATCTGCCAGCAAGGGAT 857

Qy 467 GlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeuAspIleile 486
 Db 858 GGAATGTGGATGTAATATTTATCAACATATTTTTCACATGATCTGTTTGGATATATTT 917
 Qy 487 LeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSerPheAspAla 506
 Db 918 TTAATAAACTGTTTGTAGAAAATCTGGAGAGGAGATAGTGTTCCTCATTTGATGCA 977
 Qy 507 AspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeuThrGln 526
 Db 978 GATATTTGCACAAATGTTTCGCAAAAGCAGAAACAAATATCCGATCTATTTTAACTCAA 1037
 Qy 527 GlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrProIleAla 546
 Db 1038 GGAAATCTGAGATTTATCTCTGAACTCATGGACCTCAGATCTCGGACAAACCCCATTTGCA 1097
 Qy 547 MetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGluAspLeuLeu 566
 Db 1098 ATGAGCTTTGCACAGTTTGAATAATCTACTGGGGATAAATGACATCTGAAAGACTTGCTC 1157
 Qy 567 ArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpGly 586
 Db 1158 AGAAACCCATCTATATTCAGAGGCGAAAGCTAAGGGACTAGTCATATTTCTGTGGGT 1217
 Qy 587 AspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyValAsnGlyLeu 606
 Db 1218 GATGATACCAATGATCTCTGAAACAGAGAAATTTGAAGGAACCTTGGAGTTTAATGGTCTA 1277
 Qy 607 IleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePheGlnValGluGln 626
 Db 1278 ATTTATGATAGATATATGATTTGGATGCTGAAACCAACCAATATATTTCCAGTGGAGCA 1337
 Qy 627 LeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCysLeuCysProThrValSer 646
 Db 1338 TTGGAAACGCTGACAGAGAAATGCGAGCTTAAGAGCTGTTTGTGCTCCCACTGTTAGC 1397
 Qy 647 ArgPheValProSerSerLeuCysGlyGluSerAspIleHisValAspAlaAsnGlyIle 666
 Db 1398 CGCTTTGTTCCCTCATCTTTGTGTGGGAGTCTGATATCCATGTGGATGCCAACGGCAT 1457
 Qy 667 AspAsnValGluAsnAla 672
 Db 1458 GATAACGTGGAGATGCT 1475
 RESULT 14
 ABA09583
 ID ABA09583 standard; DNA; 1978 BP.
 XX
 AC ABA09583;
 DT 15-JAN-2002 (first entry)
 DE Human bone marrow expressed oligonucleotide SEQ ID NO: 9.
 XX Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnery;
 KW antinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
 KW antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
 KW antitumor; fungicide; antidiabetic; antisthmatic; antiallergic;
 KW immunostimulant; analgesic; cerebroprotective; inflammation; infection;
 KW nervous system disorder; autoimmune disorder; allergy; ds.
 XX Homo sapiens.
 OS
 XX WO200174836-A1.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US010472.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 PR 30-NOV-2000; 2000US-0250583P.

XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT, Ford JE, Boyle BJ;
 PI WPI; 2001-626375/72.
 DR P-PSDB; ABB12339.
 DR
 PT New bone marrow-expressed nucleic acids and polypeptides, useful for
 PT diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
 PT and increasing hematopoiesis, stem cell survival and bone growth and
 PT remodeling.
 XX Claim 1; Page 163; 380pp; English.
 PS The present invention relates to bone marrow expressed polynucleotides
 XX and proteins. These sequences can be used in the treatment of
 CC inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
 CC and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
 CC disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
 CC cell disorders, platelet disorders, stem cell disorders, bone
 CC degenerative disorders, autoimmune disorders, for example multiple
 CC sclerosis, diabetes and arthritis, viral and bacterial infections,
 CC allergies and blood coagulation disorders. The present sequence is a DNA
 CC of the invention
 XX Sequence 1978 BP; 633 A; 329 C; 417 G; 599 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 3,25e-238 Length: 1978
 Score: 2439.50 Matches: 489
 Percent Similarity: 94.84% Conservative: 7
 Best Local Similarity: 93.50% Mismatches: 18
 Query Match: 68.97% Indels: 9
 DB: 5 Gaps: 3
 US-10-047-855-3 (1-672) x ABA09583 (1-1978)
 Qy 102 LeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGluThr 121
 Db 339 CTAGAAAGCGAAATTAATTTAGCATGGAGCAATTTGGAAATCCCAATGTTGTAAC 398
 Qy 122 LeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerGlu 141
 Db 399 CTGGATCTGGATGGCTGACATGTCAGACTGAATAGATTTACGTTTGCATTATCTCAA 458
 Qy 142 LysProValSerIleThrLysLysLysLysSerArgPheArgValLysLeu 161
 Db 459 AAACCTCTGTGTCATTAACCAAGAAAAATTAATAAATCTAGATTTAGGTTGAAGCTG 518
 Qy 162 ThrLeuGluGlyLeuGluGluAspAspArgValSerProThrValIleHisLys 181
 Db 519 ACACCTAGAAAGCGCTGGAGGAAGATGACGATGATAGGGTATCTCCACCTGTACTCCACAAA 578
 Qy 182 MetSerAsnSerLeuGluIleSerLeuIleSerPheAsnGluPheLysCysArgHisSer 201
 Db 579 ATGTCCAATAGCTTGGAGATATCTTTAATAAGCGCAATAGATTTCAAGTCAGGCAATCA 638
 Qy 202 GlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGlnThr 221
 Db 639 CAGCCGGAGTGTGGTTATGGCTTGCAGCTGATCGTTGGACAGAGTACAGATACAGACG 698
 Qy 222 MetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGluHis 241
 Db 699 ATGGAACACAGATAACCTCGAACTAATCTTTGATTTTTTCGAAGAAGATCTCTAGTGAGC 758
 Qy 242 ---ValValGlnGlyAspAlaLeuPro-GlyHisValGlyThrAlaCys-LeuLeu---S 259
 Db 759 CGTAAAGTTACGGGGTGATGCCCTTCTGGGACATGTGGGTACAGCTTGTCTCTTATCC 818
 Qy 259 erSerThrIleAlaGluSer--GlyLysSerAlaGlyIleLeuThrLeuProIleMetSe 278

Db 819 ATCCCAACCATGCTGAAGAGTGGAAAGAGTGTCTGGAATCTTACTTCTCCCATCATGAG 878
Qy 278 rArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLe 298
Db 879 CAGAAATTCCTCCGAAACAAATAGCAAGAGGAGATTGACTATATAATTATTAAAGCCATT 938
Qy 298 uProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTTrpLysProArgIlePr 318
Db 939 ACCAGAGATACAGTTGTGACATGAATCTTCATTTTCCAAAGTATTTGAAGCCAAAGATACC 998
Qy 318 oleuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysVa 338
Db 999 ATTGGATGTGGCCATCGAGGGCAGAAACTCTACAACTGCCAGCTGCTGCTAAAGT 1058
Qy 338 lGlnGluAsnThrIleAlaSerLeuArgAsnAlaLaSerHisGlyValaAlaPheValG1 358
Db 1059 TCAAGAAATACTATTGCTTCTTTAAGAAATGCTGTAGTCATGGCGCAGCCTTTGTAGA 1118
Qy 358 upheAspValHisLeuSerLysPheValProValValTyrHisAspLeuThrCysCy 378
Db 1119 ATTTGACGTACACCTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTTG 1178
Qy 378 sLeuThrMetLysLysPheAspAlaAspProValGluLeuPheGluIleProValLy 398
Db 1179 CTTGACTATGAAGAAAGAAATTTGATGCTGATCCAGTTGAATTAATTTGAAATTCAGATAA 1238
Qy 398 sGluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSe 418
Db 1239 AGAATTAACATTTTGACCACTCCAGTTGTTAAAGCTCACTCAITGCTGACTGCACCTGAAATC 1298
Qy 418 rLysAspArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPh 438
Db 1299 TAAGGATCGAAAGAAATCTGTGTTTCAGGAGGAAATTCCTTTTCAGAAATCAGCCATT 1358
Qy 438 eProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluI 458
Db 1359 TCCTTCTCTTAAGATGGTTTTAGAGTCTTTTGCAGAGAGATGTAGGGTTTAAACATTGAAAT 1418
Qy 458 eLysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAs 478
Db 1419 AAATCGATCTGCCAGCAAGGATGGATGTGGATGTAATTAATCAACATATTTTGA 1478
Qy 478 pMetAsnLeuPheLeuAspIleLeuLysThrValLeuGluAsnSerGlyLysArgAr 498
Db 1479 CATGAATCTGTTTTTGATATATAATTTTAAAGATCTTTTAAAGAAATCTCGGAAGAGGAG 1538
Qy 498 gIleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLy 518
Db 1539 AATAGTGTGTTTTCTTCATTTGATCGATATTTGCACAATGGTTTCGCAAAAGCAGAACAA 1598
Qy 518 sTyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspL 538
Db 1599 ATATCCGATACATATTTTAACTCAGAAATCTGTAGATTGATGCCAGTGCATGACT 1658
Qy 538 euArgSerArgThrThrProIleAlaMetSerPheAla-GlnPheGluAsnLeuLeuGly 557
Db 1659 TCAGGTGTGCGCAACACCCCATTCGAATGAGCTTTGCACGAGTTTGAAATCTGCTCGCG 1718
Qy 558 lIleAsnValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAla 577
Db 1719 ATAAATGTACATCTGAAGACTTGTCTCAGAAACCCATCTATATTCAGAGGCGAAAGCT 1778
Qy 578 LysGlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLys 597
Db 1779 AAGGGACTAGTCATATCTCTGGGTGATGATACCAATGATCTGAAAACAGAGAGAA 1838
Qy 598 Leu---LysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMet 615
Db 1839 ATTGAAGGAAACTTCGAGTTAATGCTCTAATTTATGATAGGTATTTGTTTTCATG 1895

RESULT 15
ID ADJ57947
ADJ57947 standard; cDNA; 2393 BP.

Alignment Scores: 6.14e-132 Length: 2393
Pred. No.: 1399.00 Matches: 261
Score:

XX AC ADJ57947;
XX DT 06-MAY-2004 (first entry)
XX DE Rat NARC 9 cDNA.
XX KW Neuronal apoptosis regulated candidate; NARC; diagnosis; cancer;
KW lung disease; cirrhosis; hepatitis; inflammation; atherosclerosis;
KW myocardial infarction; inflammation; anaemia; glomerulonephritis;
KW osteoporosis; AIDS; acquired immunodeficiency syndrome;
KW Parkinson's disease; Alzheimer's disease; stroke; dermatitis;
KW drug screening; gene therapy; cytostatic; hepatotropic; nootropic;
KW cerebroprotective; dermatological; virucide; neuroprotective;
KW phosphatidylglycerolphosphate synthase; FCP synthase; rat; ss.
XX OS Rattus norvegicus.
XX PN US2004009553-A1.
XX PD 15-JAN-2004.
XX PF 30-APR-2003; 2003US-00426776.
XX PR 27-SEP-1999; 99US-00406045.
XX PR 22-OCT-1999; 99US-0161188P.
XX PR 31-JAN-2000; 2000US-00495823.
XX PR 28-FEB-2000; 2000US-0185517P.
XX PR 20-OCT-2000; 2000US-00692785.
XX PR 31-JAN-2001; 2001US-00773426.
XX PR 28-FEB-2001; 2001US-00795691.
XX PR 31-OCT-2001; 2001US-0335003P.
XX PR 31-OCT-2001; 2001US-0335037P.
XX PR 25-MAR-2002; 2002US-00105992.
XX PR 28-AUG-2002; 2002US-00229662.
XX PR 30-OCT-2002; 2002US-00284014.
XX PR 30-OCT-2002; 2002US-00284059.
XX PR 09-DEC-2002; 2002US-00314881.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Glucksmann MA, Williamson MJ, Tsai F, Rudolph-Owen LA;
PI Kapeller-Libermann R, Meyers RE, Chiang LW, Hunter JJ;
XX WPI; 2004-090469/09.
XX PT New isolated nucleic acid molecules and polypeptides (e.g. 27411, 23413,
PT 22438, 23553, NARC SC1 or NARC 1) useful for diagnosing, preventing or
PT treating disorders associated with the protein, e.g. cancer,
PT atherosclerosis or AIDS.
XX PS Claim 1; SEQ ID NO 48; 260pp; English.
XX CC The present invention provides isolated nucleic acid molecules and
CC proteins designated 27411, 23413, 22438, 23553, 25278, 23412, NARC SC1,
CC NARC 10A, NARC 1, NARC 12, NARC 13, NARC 17, NARC 25, NARC 3, NARC 4,
CC NARC 7, NARC 8, NARC 11, NARC 14A, NARC 15, NARC 16, NARC 19, NARC 20,
CC NARC 26, NARC 27, NARC 28, NARC 30, NARC 3, NARC 9, NARC 10C,
CC NARC 8B, NARC 9, NARC 2A, NARC 16B, NARC 1C, NARC 1A, NARC 25, 86604 or
CC 32222. The invention is useful in diagnosing, preventing or treating
CC disorders such as cancer, lung diseases, cirrhosis, hepatitis,
CC atherosclerosis, myocardial infarction, inflammation, anaemia,
CC glomerulonephritis, osteoporosis, AIDS (acquired immunodeficiency
CC syndrome), Parkinson's disease, Alzheimer's disease, stroke and
CC dermatitis. These may also be used in drug screening. The invention is
CC also useful in gene therapy. The present sequence is rat neuronal
CC apoptosis regulated candidate (NARC) cDNA.
XX SQ Sequence 2393 BP; 762 A; 435 C; 455 G; 741 T; 0 U; 0 Other;

Tue Jul 5 14:07:49 2005

Percent Similarity: 95.85% Conservatives: 16
 Best Local Similarity: 90.31% Mismatches: 12
 Query Match: 39.55% Indels: 0
 DB: 12 Gaps: 0

US-10-047-855-3 (1-672) x ADJ57947 (1-2393)

Qy	384	LysPheAspAlaAspProValGluLeuPheGluLeuProValLysGluLeuThrPheAsp	403
Db	71	AAATATGAAGCTGATCGAGTTGAATTTGTTGAAATCCAGTAAAGGAATTAACATTCGAC	130
Qy	404	GlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGlu	423
Db	131	CAACTCCAGTTATTGAAGCTTTCATGTGACTGCACCTAAACCAAGACCAAGACCAACAA	190
Qy	424	SerValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMet	443
Db	191	TGTATGGCTGAGGAGGANAATTCCTTTCTGAAACCAACCATTTCTCTTCTTAAGATG	250
Qy	444	ValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGln	463
Db	251	GTTTTAGAGTCATTGCCAGAAAATGTAGGATTTAATATAGAAATAAAATGGATTTGCCAA	310
Qy	464	GlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeu	483
Db	311	CACAGGGATGGAGTAGTGGAGCGCAACTTATCGACATATTTTGTATGAAATGCATTTTGG	370
Qy	484	AspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSer	503
Db	371	GATATPAAATTTTAAAACCTGTTTGTAGAAAATTCGGGAAGAGAGAGAAATAGTATTTCTTCA	430
Qy	504	PheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPhe	523
Db	431	TTTGTATGCAGACATCTGTACATGTTGGCAGAGAACAAACAAATATCCCATATTTATTT	490
Qy	524	LeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrThr	543
Db	491	TTGACCCCAAGGAAGTCTGACATTTTACCTGAACTCATGGACCTCAGATCTCGGACAAACA	550
Qy	544	ProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHisThrGlu	563
Db	551	CCCATTTGCAATGAGCTTTGCACAGTTTGAANAATATTTGGGGGATAAATGCCCATACTGAA	610
Qy	564	AspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePhe	583
Db	611	GATCTCCTTTAGAAACCCATCTATGTCTCAAGAGGCAAAAGATAAGGGATTTGGTCATATTC	670
Qy	584	CysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyVal	603
Db	671	TGCTGGGGTGATGATACCAATGATCTGAAAACAGAAAGGAACTGAAGGAATTTGGAGTA	730
Qy	604	AsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePheGln	623
Db	731	AATGGTCTAATATATATAGATAGATATACGATTGGATCGCTGACACCAACAAATATATCCAA	790
Qy	624	ValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCysLeuCysPro	643
Db	791	GTGGAGCAACTGGAGCGCTGGAAGCGAGAATTCACAGAGCTTAAGAACTGTTTGTGTCCC	850
Qy	644	ThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHisValAspAla	663
Db	851	ACTGTTAGCCACTTCAATCTCTCTTCTTCTGTATGGAGTCTAAATATCCATGTGGATGCT	910
Qy	664	AsnGlyIleAspAsnValGluAsnAla	672
Db	911	AACGGCATTGATATGTGGAGACGCT	937

Search completed: July 3, 2005, 07:07:33
 Job time : 1016 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 3, 2005, 06:45:04 ; Search time 8001 Seconds
(without alignments)
4069.729 Million cell updates/sec

Title: US-10-047-855-3
Perfect score: 3537
Sequence: 1 MTPSQVAFIRGTLPLGEVP.....LCGESDIHVANDINDVENA 672

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2_1/USPTO_spool/US10047855/runat_01072005_155528_2220/app_query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-INIT=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10047855 @CNG 1.1 5081 @runat 01072005_155528_2220 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3537	100.0	3206	6	AX127528 Sequence
2	3537	100.0	3206	6	AX675186 Sequence
3	3537	100.0	3265	9	BC027588 Homo sapi
4	3537	100.0	3499	6	CQ413929 Sequence

5	3537	100.0	3499	6	CQ491863 Sequence
6	3537	100.0	3499	6	CQ497765 Sequence
7	3537	100.0	5443	9	AB037855 Homo sapi
8	3254	92.0	3229	10	AY233980 Rattus no
9	2992.5	84.6	39694	2	AC144439 Homo sapi
10	2992.5	84.6	149259	9	AC099654 Homo sapi
11	2979	84.2	6571	9	AL833069 Homo sapi
12	2976.5	84.2	160762	2	AC150686 Pan trogl
13	2938	83.1	1731	6	CQ728047 Sequence
14	2897.5	81.9	3174	10	BC033408 Mus muscu
15	2552	72.2	2738	6	AX127525 Sequence
16	1364	38.6	1803	6	BD156897 Primer fo
17	1364	38.6	1803	6	AX877873 Sequence
18	1364	38.6	1803	6	AX001947 Homo sapi
19	1331.5	37.6	4353	10	AK122510 Mus muscu
20	1102.5	31.2	768	6	BD149663 Primer fo
21	1102.5	31.2	768	6	AX869601 Sequence
22	1041	29.4	85916	2	AC023409 Homo sapi
23	1005	28.4	2774	3	AY069505 Drosophil
24	1005	28.4	2792	6	CQ577356 Sequence
25	922	26.1	5439	6	CQ577355 Sequence
26	922	26.1	46847	2	AC019993 Drosophil
27	922	26.1	98942	3	AC005149 Drosophil
28	922	26.1	155566	3	AC092231 Drosophil
29	922	26.1	187707	3	AC092240 Drosophil
30	922	26.1	304674	3	AE003579 Drosophil
31	851	24.1	245453	2	AC131004 Rattus no
32	851	24.1	252563	2	AC094054 Rattus no
33	843	23.8	1999	3	AK112176 Ciona int
34	717	20.3	520	6	CQ409598 Sequence
35	667	18.9	2392	6	CQ599424 Sequence
36	667	18.9	2457	6	CQ599139 Sequence
37	667	18.9	3013	3	AY061498 Drosophil
38	667	18.9	7162	6	CQ599138 Sequence
39	667	18.9	9667	6	CQ599423 Sequence
40	667	18.9	174054	2	AC015160 Drosophil
41	667	18.9	185497	3	AC010563 Drosophil
42	667	18.9	279088	3	AE003518 Drosophil
43	666	18.8	2299	6	CQ592209 Sequence
44	666	18.8	2329	3	AY070640 Drosophil
45	666	18.8	4299	6	CQ592208 Sequence

ALIGNMENTS

RESULT 1	AX127528	Sequence 4 from Patent WO0131007.	3206 bp	DNA	linear	PAT 15-MAY-2001
LOCUS	AX127528	Sequence 4 from Patent WO0131007.				
DEFINITION	AX127528					
ACCESSION	AX127528					
VERSION	AX127528.1	GI:14134233				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1					
AUTHORS	Chiang L.W.					
TITLE	Nucleic acid molecules derived from rat brain and programmed cell death models					
JOURNAL	Patent: WO 0131007-A 4 03-MAY-2001;					
FEATURES	Millennium Pharmaceuticals, Inc. (US)					
source	Location/Qualifiers					
	1..3206					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					

ORIGIN						
Alignment Scores:	8.76e-298	Length:	3206			
Pred. No.:	3537.00	Matches:	672			
Score:	100.00%	Conservative:	0			

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-047-855-3 (1-672) x AX127528 (1-3206)			
Qy	1	MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe	20
Db	145	ATGACACCTTCTCAGGTGGCTTTGAAATATAGAGGAATCTTTTACGAGGAGAAGTTTT	204
Qy	21	AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu	40
Db	205	CGGATATGTGAAGCTGTGATGCTTTGGGAACTGGGAATCTCTCAAAATGCTGTGGCTCTT	264
Qy	41	LeuProGluAenAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg	60
Db	265	CTTCCAGAGATGACACAGGTGAAAGCATGTATGGAAAGCAACATTTGACTCAGTAGA	324
Qy	61	GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly	80
Db	325	GGAGTATCAGTTTCAGTATCGTACTTCAAGGGTACTTTTATAGAACCAAGACATTCGGT	384
Qy	81	GlyProCysGlnValIleValHisIysTrpGluThrHisLeuGlnProArgSerIleThr	100
Db	385	GGTCATGTCAGGTGATAGTTTCAAGTGGGAGACTCATCTACAAACCAAGCATCAATACC	444
Qy	101	ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu	120
Db	445	CCTTTAGAAAGCGAAATTTATTGACGATGGACAAATTTGGAATCCCAATGGTGTGAA	504
Qy	121	ThrLeuAspSerGlyTyrLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer	140
Db	505	ACTCTGGATTTCTGGATGCTGACATGTCAGACTGAAATAGAAATACGTTTGCATTTCT	564
Qy	141	GluLysProProValSerIleThrLysLysLysLysLysSerArgPheArgValLys	160
Db	565	GAATAACCTCTGCTGTCATTAACCAAGAAATAATTAATAAATACTTAGAATTTAGGGTGAG	624
Qy	161	LeuThrLeuGluGlyLeuGluAspAspArgValSerProThrValLeuHis	180
Db	625	CTGACACTTAGAAGGCTGGAGGAGATGACGATGATAGGGTATCTCCACTGTACTCCAC	684
Qy	181	LysMetSerAsnSerLeuGluIleSerLeuLeuSerAspAsnGluPheLysCysArgHis	200
Db	685	AAATGTTCCATAGCTTGGAGATATCTTAATAGCGACAATGAGTCAAGTGCAGGCAT	744
Qy	201	SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTyrThrGluTyrSerIleGln	220
Db	745	TCACAGCCGGAGTGGTGTATGGCTTGCAGCCCTGATCGTTGGACAGAGTACAGACATACAG	804
Qy	221	ThrMetGluProAspAsnLeuGluLeuIlePheAspPhePheGluGluAspLeuSerGlu	240
Db	805	ACGATGGAAACCGAGATAACCTGGAACTATCTTTTGAATTTTTCGAAGAGAGATCTCAGTGAG	864
Qy	241	HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer	260
Db	865	CACGTAGTTCAGGGTATGCCCTTCTCGACATGTGGGTACAGCTTGTCTCTTATCATCC	924
Qy	261	ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn	280
Db	925	ACCAATGCTGAGAGTGGAAAGAGTGTGGAAATCTTTACTCTTCCATCATAGACAGAAAT	984
Qy	281	SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly	300
Db	985	TCCCGGAAACCAATAGGCAAGTGAGAGTTGACTATATTAATTTATTAAGCCATTTACAGGA	1044
Qy	301	TyrSerCysAspMetLysSerPheSerLysTyrTrpLysProArgIleProLeuAsp	320
Db	1045	TACAGTTGTGACATGAAATCTTCATTTTCCAGTATTTGGAAGCCAAAGATACCATTTGGAT	1104
Qy	321	ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu	340
Db	1105	GTTGGCCATCGAGGTGCAGGAACTCTACAACAACCTGCCAGCTGGCTAAAGTTCAAGAA	1164
Qy	341	AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp	360
Db	1165	AATACATATTGCTTCTTTTAAAGAAATGCTGTAGTCATGGTGCAGCCCTTTGTAGAAATTTGAC	1224
Qy	361	ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr	380
Db	1225	GTACACCTTTCAAGGACCTTTGTGCCGTGGTATATCATGATCTTACCTGTGTGTGACT	1284
Qy	381	MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu	400
Db	1285	ATGAAAAAGAAATTTGATGCTGATCCAGTTGAAATTTTCAAAATTCAGTAAAAAGAAATTA	1344
Qy	401	ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp	420
Db	1345	ACATTTGACCACTCCAGTTGTTAAGCTCCTCATGTGACTGCACCTGAAATCTAAGGAT	1404
Qy	421	ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer	440
Db	1405	CGGAAAGAATCTGTGGTTTCAGGAGGAAAAATTCCTTTTCAGAAAAATCAGCCATTTCTTCT	1464
Qy	441	LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp	460
Db	1465	CTTAAGATGCTTTTAGAGTCTTTGCCAGAAAGATGTAGGGTTTAAACATTTGAATAAAAATGG	1524
Qy	461	IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn	480
Db	1525	ATCTGCCAGCAAGGGATGGAAATGTGGGATGGTAACCTTATCAACATATTTTTCACATGANT	1584
Qy	481	LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal	500
Db	1585	CTGTTTTTGGATATATATTTTAAAAAAGTCTTTTAAAAAATCTCGGAAAGAGAGAAATAGTG	1644
Qy	501	PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro	520
Db	1645	TTTTCTTCAATTTGATGACAGATATTTGCAATGTTTCGGCAAAAGCAGAAACAAATATCCG	1704
Qy	521	IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer	540
Db	1705	ATATCTATTTTAACTCAAGGAAATCTGAGATTTATCTCTGAACCTCATGGACCTCAGATCT	1764
Qy	541	ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal	560
Db	1765	CGGCAACCCCACTGCATGAGCTTTGACAGCTTTGAAATCTACTGGGATATAATGTA	1824
Qy	561	HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu	580
Db	1825	CATACCTGAAGACTTGTCTCAGAAACCCATCTATATTTCAAGAGGCAAAAGCTTAAGGACTA	1884
Qy	581	ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu	600
Db	1885	GTCTATTTCTCTGGGGTGATGATACCAATGATCTCTGAAAAAGAGAAATTTGAAGAA	1944
Qy	601	LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn	620
Db	1945	CTTGAGTTAATGGTCTAATTTTATGATAGGATATATGATTTGGATGCTTGAACACCAAT	2004
Qy	621	IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys	640
Db	2005	ATATTTCAAGTGGAGCAATTCGAACGCTGAAGCAGGAAATTCAGAGAGCTTAAGAGCTGT	2064
Qy	641	LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis	660
Db	2065	TTGTGCTCCACTGTTAGCGCTTTGTTCCCTCATCTTTGTGTGGGGAGTCTGATATCCAT	2124
Qy	661	ValAspAlaAsnGlyIleAspAsnValGluAsnAla	672
Db	2125	GTGGATGCCAACCGCAATTGATAACGTGGAGAAATGCT	2160
RESULT 2			
AX675186			
LOCUS			
DEFINITION Sequence 4 from Patent WO02081516.			
AX675186 3206 bp DNA linear PAT 27-MAR-2003			

```
ACCESSION AX675186
VERSION AX675186.1 GI:29333417
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Chiang L.W.
Narc10 and narc16, programmed cell death-associated molecules and
uses thereof
Patent: WO 02081516-A 4 17-OCT-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. 3206
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature
1..17
/note="Vector sequence"
CDS
145..2163
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD83808.1"
/db_xref="GI:29333418"
translation="MTPSQVAFEIRGTLLPGSEVFALCGSCDNLNPNQNAVALLPEN
DTGESMLKATIVLPGSVQVYFKGYFLEPKTIGGPGQVIVHKWETHLOPRISITPL
ESEIIDDQGFHNGVETLDSGMLTCOTEIRLRLHYSEKPPVITKKLKKSRFRVK
LTLEGLDEDDRVSPVTLHKMSLEISLISDNFEKCHSOPGCGYGLQDPRTWEYS
IQTMPDNLELIFDFEEDLSHVQGDALPGHVTACLLSSTIAESKSGAGILILPI
MERNRKTITGRVVDYIIIKPLPGYSCDMKSSFSKYWKPRIPLDVGRHAGNSITTAQ
LAKQENTIASLRNAAASHGAFFVDFVHLSDKDFVPPVYHDLTCLTMKKKFDADPEL
FEIPVKEITFDQLQLKLTHTVALKSKDKESVQVQENSFSENQPPSLKMYLESLEPE
DVFNEIKWICQORDGMDGNLSTYFDMNLFLDIILKTVLENSGKRRIVFSSFDADI
CTMVKQKNKYPIILFTOCKSRIYELMDLRSRTTPIAMSAFQENLGLINVTEDLL
RNPSTIQEKAKGLVIFCMGDDTNDPNERKKLKELVNGLIYDRYIDWNPQEPNIFQV
EQLERKLQELPELKSCLCPTVSRFPVSSLCGESDIHVDANGIDNVENA"
ORIGIN
Alignment Scores:
Pred. No.: 8,76e-298 Length: 3206
Score: 3337.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-047-855-3 (1-672) x AX675186 (1-3206)
QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 145 ATGACACCTTCTCAGGTTGCCCTTTGAAATAAGAGGAACCTCTTTTACCAGGAGAGTTT 204
QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 205 GCGATATGTGGAGCTGTGATGCTTTGGGAAACCTGGAATCCTCAAAATGCTGTGGCTCT 264
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 265 CTTCCAGAGAATGACACAGTGAACATGCTATGGAAGCAACCACTTGTACTCAGTAGA 324
QY 61 GlyValSerValGlnIleArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 325 GGAGTATCAGTTTCAGTATCGCTACTTCAAGGGTACTTTTAGAACCAAGACTATCGGT 384
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 385 GGTCCATGTCAAGTGATAGTTCACAAAGTGGGAGCTCATCTACAACCAAGCATCAATACC 444
QY 101 ProLeuGluSerGluIleIleLeuAspGlyGlnPheGlyIleHisGlnGlyValGlu 120
Db 445 CCTTTAGAAAGCGAAATTATTATTGACGATGACAAATTTGGAATCCACAATGGTGTGAA 504
```

```
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
Db 505 ACTCTGGATTCTGGATGGCTGACATGTTCAGACTGAAATAAGATTACGTTTGCATTATCT 564
QY 141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
Db 565 GAAAACCTCTCTGTGTCAATAACCAAGAAAAAATTAATAAATCTAGATTGGGTGAAG 624
QY 161 LeuThrLeuGluGlyLeuGluAspAspArgValSerProThrValLeuHis 180
Db 625 CTGACACTAGAAAGCGCTGGAGGAAGATGACGATAGGGTATCTCCCACTGTACTCCAC 684
QY 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
Db 685 AAAATGTCCAATAGCTTGGAGATATCTTTAATAAGCGACAATGAGTCAAGTCAGGCAT 744
QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
Db 745 TCACAGCCGGAGTGTGTATGGCTTGCAGCTGTATCGTTGGACAGAGTACAGCATACAG 804
QY 221 ThrMetGluProAspAsnLeuGluLeuPheAspPheGluGluAspLeuSerGlu 240
Db 805 ACGATGGAAACACAGATAACCTGGAACATACTTTTTCGAAGAAGATCTCAGTGAG 864
QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db 865 CACGTAGTTTCAGGGTGATGCCCTTCTCTGGACATGTGGGTACAGCTTGTCTTATCATCC 924
QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
Db 925 ACCATTGCTAGAGTGGAAAGAGTGTGGAATCTTACTTCTTCCCATCATGAGCAAAAT 984
QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300
Db 985 TCCCGGAAAAACAATAGGCAAGTGTGAGATGATATATAATTAATTAAGCCATTACCAGGA 1044
QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
Db 1045 TACAGTTGTGACATGAAATCTTCAATTTCCAAAGTATTGGAAGCCAAAGATACCATTTGGAT 1104
QY 321 ValGlyHisArgGlyValAlaGlyAsnSerThrThrAlaGluLeuAlaLysValGlnGlu 340
Db 1105 GTTGCCCATCGAGGTGCGAGAAACTCTACAACACTGCCAGCTGGCTAAGTTCAAGAA 1164
QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyValAlaPheValGluPheAsp 360
Db 1165 AATACTATTGCTTCTTTAAGAAATGCTGCTAGTCAATGGTGCAGCCCTTTGTAGAAATTGAC 1224
QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
Db 1225 GTACACCTTTCAAAGGACTTTTGTGCCGCTGTATATCATGATCTTACCTGTGTTGACT 1284
QY 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
Db 1285 ATGAAAAAGAAATTTGATGCTCATCCAGTTGAATATTATTGAAATTCAGATAAAGAAATTA 1344
QY 401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
Db 1345 ACATTTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCATGAAATCTAAGGAT 1404
QY 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
Db 1405 CGGAAGAATCTGTGTTTCAGGAGGAAATTCCTTTTCAGAAAAATCAGGCATTTCTCTCT 1464
QY 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluLysTrp 460
Db 1465 CTTAAGATGGTTTATAGAGTCTTTGCGAGAAAGATGAGGGTTTAAACATGAAATAAATGG 1524
QY 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
Db 1525 ATCTCCAGCAAGGAGTGAATGTGGATGTGTAATCTATCAACATATTTTTCACATGAAT 1584
QY 481 LeuPheLeuAspIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal 500
```


Db	1585	CTGTTTGGATATAATTTTAAACCTGTTTGAATAATCTGGGAAGAGGAAATAGTG	1644		
Qy	501	PheSerSerPheAspAlaAaspIleCysThrMetValArgGlnLysGlnAenLysTyPro	520		
Db	1645	TTTTTCTTCATTTGATGAGATATTTTGCAATGTTTGGCAAGACAGACAAATATCCG	1704		
Qy	521	IleLeuPheLeuThrGlnGlyLysSerGluIleTyProGluLeuMetAspLeuArgSer	540		
Db	1705	ATACTATTTTAACTCAAGGAAATCTGAGATTATCTGAACTCATGGACCTCAGATCT	1764		
Qy	541	ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAenLeuGlyIleAenVal	560		
Db	1765	CGGACAAACCCCATTTGCAATGAGCTTTTCACAGTTTGAAAATCTACTGGGGATAAATGTA	1824		
Qy	561	HisThrGluAepLeuLeuAqAenProSerTyrtileGlnGluAlaLysAlaLysGlyLeu	580		
Db	1825	CATACTGAAGACTTGCTCAGAAACCCATCTATATTCAAGAGGCAAAAGCTAAGGACTA	1884		
Qy	581	ValIlePheCysTrpGlyAaspThrAsnAaspProGluAenArgArgLysLeuLysGlu	600		
Db	1885	GTCAATTTCTCTGGGTGATGATACCAATGATCTCTGAAAACAGAGAAATTCAGAGAA	1944		
Qy	601	LeuGlyValAenGlyLeuIleTyAspArgIleTyAspTrpMetProGluGlnProAen	620		
Db	1945	CTTGAGTTAATGGTCTAAATTTATGATAGGATATATGATGGATGCCCTGAACACCAAT	2004		
Qy	621	IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys	640		
Db	2005	ATATTCCAAGTGGAGCAATTTGGAACGCTGGAAGCAGGAATTTGCCAGACCTTAAGAGTGT	2064		
Qy	641	LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGlyUserAspIleHis	660		
Db	2065	TTGTGTCCTCACTGTAGCGCTTTGTTCCCTCACTTTGTGTGGGAGTCTGATATCCAT	2124		
Qy	661	ValAspAlaAenGlyIleAaspAenValGluAenAla	672		
Db	2125	GTGGATGCCAACGGCATTTGATAACCTGGAGAAATGCT	2160		
RESULT 3	BC027588	3265 bp	mRNA	linear	PRI 30-JUN-2004
LOCUS	BC027588	Homo sapiens hypothetical protein KIAA1434, mRNA (cDNA clone			
DEFINITION	MGC:26147 IMAGE:4827346), complete cds.				
ACCESSION	BC027588				
VERSION	BC027588.2	GI:34194199			
KEYWORDS	MGC.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	1 (bases 1 to 3265)				
	Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Hunkeler,D., Venter,A., Auletta,D., Rozen,S., Hornik,M., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stabile,C.M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				

12477932	PUBMED			
2 (bases 1 to 3265)	REFERENCE			
Strausberg,R.	AUTHORS			
Direct Submission	TITLE			
Submitted (08-APR-2002) National Institutes of Health, Mammalian	JOURNAL			
Gene Collection (MGC), Cancer Genomics Office, National Cancer				
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
USA				
NIH-MGC project URL: http://mgc.nci.nih.gov	REMARK			
On Aug 25, 2003 this sequence version replaced gi:20379525.	COMMENT			
Contact: MGC help desk				
Email: cgaphs-r@mail.nih.gov				
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.				
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki				
Toshiyuki and Piero Carninci (RIKEN)				
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
DNA Sequencing by: Institute for Systems Biology				
http://www.systemsbio.org				
contact: amadan@systemsbiology.org				
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha				
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting				
Clone distribution: MGC clone distribution information can be found				
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov				
Series: IRAK Plate: 34 Row: a Column: 18				
This clone was selected for full length sequencing because it				
passed the following selection criteria: matched mRNA gi: 25901061.				
Location/Qualifiers	FEATURES			
1. .3265	source			
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="MGC:26147 IMAGE:4827346"				
/tissue_type="Testis"				
/clone_lib="NIH_MGC_97"				
/lab_host="DH10B"				
/notes="Vector: pBluescript"				
1. .3265	gene			
/gene="KIAA1434"				
/note="synonyms: FLJ11085, MGC26147"				
/db_xref="LocusID:56261"				
213. .2231	CDS			
/gene="KIAA1434"				
/codon_start=1				
/product="hypothetical protein KIAA1434"				
/protein_id="AAH27588.1"				
/db_xref="GI:20379526"				
/db_xref="LocusID:56261"				
/translation="MTPSQVAFRTGTLTPGEVFAICGSDALGNWNPONAVALLPEN				
DTGESLWUKATIVLSRGVSQYRYFKYGFLEPKTIGPCQVIVHKWETHLPQSRITPL				
ESIIIDDGFGIHNGVETLSDGWLTCQTEIRLKLHYSEKPPVSTIKKKLKKSRFRK				
LTLEGEEDDDRVSTPLVHKMSNSLEISLSDNEFKRSQCEGGLQPDWRVTS				
IQTMENLELIDFPEEDLSEHVQGDALPGHVTACLLSSTIAESGKSAGILTLPI				
MSRNRKTIQKRVADYIIIKPLPGYSCDKMSKSKYWKPRIPLDVGHGAGNSTTQ				
LAKVOENTSLRNAASHGAAFEFVHLSKDFVPVYVHDLTCCLTWKKFADAPVEL				
FEIPKSLTFDQQLKLHTVALKSKDKESVVOENSPSENQPPSLKXVLSLPE				
DVGNIEIKWICQQRDGMGNLSYTFDMNLFLDIILKTVAENSGKRIIVSFSDADI				
CTVVRQKQNPILFTQKGSSEIYFELMDLRSRTPTIAMSFAQENLLGNVHEDLL				
RNPSYIOEAKAGLIVFCWGDNDNPNRKLKELGVNGLIYDRIYDMVQPNIFQV				
EQLERLQELPELKSCLCTPVSFVPSLCSGESDIHVDANGIDNVENA"				
Alignment Scores:				
Pred. No.: 8.98e-298	Length:	3265		
Score: 3537.00	Matches:	672		
Percent Similarity: 100.00%	Conservative:	0		
Best Local Similarity: 100.00%	Mismatches:	0		
Query Match: 100.00%	Indels:	0		
DB: 9	Gaps:	0		
US-10-047-855-3 (1-672) x BC027588 (1-3265)				
Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20				

|||||
213 ATGACACCTTCTCAGGTGCTTTGAAATAAGAGGAACCTTTTACAGGAGAAGTTT 272
Qy
21 AlaileCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db
273 GCGATATGTGGAAGCTGTGATGCTTTGGGAAACTGGAATCCTCAAAATGCTGTGGCTCT 332
Qy
41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpIlyAlaThrIleValLeuSerArg 60
Db
333 CTTCCAGAGAATGACACAGGTGAAAGCATGCTATGGAAGCAACCACTGTACTCAGTAGA 392
Qy
61 GlyValSerValGlnTrpArgTrpPheLysGlyTrpPheLeuGluProLysThrIleGly 80
Db
393 GAGTATCAGTTCTCAGTATCGTACTTCAAGGGTACTTTTTAGAACCAAGACTATCGGT 452
Qy
81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db
453 GGTCCATGTCAAGTGATGTTCAACAAGTGGGACATCATACCAACCAAGCAATAACC 512
Qy
101 ProLeuGluSerGluIleIleLeuAspAspGlyGlnPheGlyIleHisGlyValGlu 120
Db
513 CCTTTAGAAAGCGAAATTTATTTAGCATGGACAATTTGGAATCCACAATGGTGTGAA 572
Qy
121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgGluLeuHisTrpSer 140
Db
573 ACTCTGGATTCCTGGATGGCTGACATGTCACTGAAATAGATTACGTTTGCATTAATCT 632
Qy
141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
Db
633 GAAAAACCTCCTGTGTCATACCAAGAAATAATTAATAAATCTAGATTAGGTGGAAG 692
Qy
161 LeuThrLeuGluGlyLeuGluGluAspAspAspArgValSerProThrValLeuHis 180
Db
693 CTGACACTAGAAGCCCTGGAGGAAGATGACGATAGGTATCTCCCACTGTACTCCAC 752
Qy
181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
Db
753 AAAATGTCGAATAGCTTGGAGATATCCTTAATAAGCGCAATGAGTTCAAGTGCAGGCAT 812
Qy
201 SerGlnProGluCysGlyTrpGlyLeuGlnProAspArgTrpThrGluTrpSerIleGln 220
Db
813 TCACAGCCGAGTGTGGTTATGCTTGCAGCCGTGATCGTTGACAGAGTACAGCATACAG 872
Qy
221 ThrMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGlu 240
Db
873 ACGATGGAACACAGATAACCTGGAACTAACTTTGATTTTTCGAAGAAGATCTCAGTGAG 932
Qy
241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db
933 CACGTAGTTTCAGGGTGATGCCCTTCTCGACATGTGGGTACAGCTTGTCTCTATCATCC 992
Qy
261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
Db
993 ACAATGCTGAGAGTGGAAAGAGTGTGGAATTTCTTACTCTTCCCATCATGAGCAAAAT 1052
Qy
281 SerArgLysThrIleGlyLysValArgValAspTrpIleIleLysProLeuProGly 300
Db
1053 TCCCGAAACCAATAGGCAAGAGTGAGAGTTGACTATATAATTAAGCCATTAACAGGA 1112
Qy
301 TyrSerCysAspMetLysSerSerPheSerLysTrpTrpLysProArgIleProLeuAsp 320
Db
1113 TACAGTTGTGACATGAATCTTTCATTTTCCAAAGTATTGGAAGCCAAAGATACCATTTGAT 1172
Qy
321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
Db
1173 GTTGGCCATCGAGGTGCAGAAACTCTCAACAACATGCCAGCTGGCTAAAGTTCAAGAA 1232
Qy
341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
Db
1233 AATACATATTGCTTCTTTAGAAATGCTGCTAGTATGCTGAGCCCTTGTAGAAATTGAC 1292
Qy
361 ValHisLeuSerLysAspPheValProValValThrHisAspLeuThrCysCysLeuThr 380
|||||

Db
1293 GTACACCTTTCAAAGCATTGTGCGCGTGGTATATCATGATCTTACCTGTTGTTGACT 1352
Qy
381 MetLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
Db
1353 ATGAAAAAGCAATTTGATGCTCATCCAGTTGAATTTATTTGAAATTCACGATAAAGAAATTA 1412
Qy
401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
Db
1413 ACATTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCATGAAATTAAGGAT 1472
Qy
421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
Db
1473 CGGAAAGAATCTGTGGTTCAGGAGGAAATTCCTTTTCAGAAATCAGGCATTTCTTCT 1532
Qy
441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
Db
1533 CTTAAGATGTTTATAGAGTCTTTGCCAGAAGATGTAGGCTTTAAACATTCAATATAAATGG 1592
Qy
461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTrpPheAspMetAsn 480
Db
1593 ATCTCCACAGCAAGGATGGAATGTGGATGGTAACTTATCAACATATTTTGACATGAAT 1652
Qy
481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal 500
Db
1653 CAGTTTGGATTAATTTTAAACTGTTTATAGAAATTCCTGGGAGAGGAGATAGTG 1712
Qy
501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTrpPro 520
Db
1713 TTTTCTTCATTTGATGCGAGATATTGCACAATGGTTCCGCAAAAGCAGAAATAATCCG 1772
Qy
521 IleLeuPheLeuThrGlnGlyLysSerGluIleTrpProGluLeuMetAspLeuArgSer 540
Db
1773 ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTCTGAACTCATGACCTCAGATCT 1832
Qy
541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
Db
1833 CGGACAAACCCCATTTGCAATGAGCTTTGCACAGTTTGAATACTCTCGGGGATAAATGTA 1892
Qy
561 HisThrGluAspLeuArgAsnProSerTrpIleGlnGluAlaLysAlaLysGlyLeu 580
Db
1893 CATACTGAAGACTTCTCAGAAACCATCTATATTCAGAGGCAAAAGCTAAGGACTTA 1952
Qy
581 ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
Db
1953 GTCATATCTCTGCGGGTATGATACCAATGATCTCTGAAACAGAAAGAAATTTGAAGGAA 2012
Qy
601 LeuGlyValAsnGlyLeuIleTrpAspArgIleTrpAspTrpMetProGluGlnProAsn 620
Db
2013 CTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATTGGATGCCCTGAAACCAAAAT 2072
Qy
621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnLeuProGluLeuLysSerCys 640
Db
2073 ATATTCGAAGTGGAGCAATTTGAAAGCTGAGCAGAGGAAATTCAGAGGCTTAAGAGCTGT 2132
Qy
641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyCysSerAspIleHis 660
Db
2133 TTGTGTCCACATGTTAGCGCTTGTCTCTCATCTTTGTGTGGGAGTCTGATATCCAT 2192
Qy
661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db
2193 GTGGATGCCAACGGCATTGATAACGTGGAGAAATGCT 2228
RESULT 4
CQ413929
LOCUS
DEFINITION
Sequence 21000 from Patent WO0170979.
CQ413929
ACCESSION
VERSION
CQ413929.1 GI:41321710
KEYWORDS
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ORGANISM

REFERENCE 1
AUTHORS Lee, J. and Lillie, J.
TITLE Genes, compositions, kits, and method for identification,
assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 21000 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1..3499
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 9,86e-298 Length: 3499
Score: 3537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-047-855-3 (1-672) x CQ413929 (1-3499)
QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
DB 221 ATGACACCTTCTCAGGTGGCTTTGAAATAAGAGGAACCTCTTTTACCGAGGAAGTGT 280
QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
DB 281 GCGATATGTGAGAGCTGTGATGCTTTGGAACTGGAACTCTCAAAATGCTGTGGCTCT 340
QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
DB 341 CTTCCAGAGAATGACACAGGTGAAAGCATGCTATGGAAGCAACCATTTGTACTCAGTAGA 400
QY 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
DB 401 GGAGTATCAGTTTCAGTATCGCTACTTCAAAGGGTACTTTTATAGAACCAAGACATATCGGT 460
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
DB 461 GGTCCATGTCAAGTATGATAGTTCAAGTGGGAGACTCATCTACACCAACGATCAATAACC 520
QY 101 ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
DB 521 CTTTATAGAAAGCAATATTATTATTCACGATGGACAATTTGGAATCCACAATGGTGTGAA 580
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
DB 581 ACTCTGGATTCTGGATGGCTGACATGTCTCAGACTGAATAAGATTACGTTTGCATTTCT 640
QY 141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
DB 641 GAAAAACCTCTCTGTGTCAATACCAAGAAAAAATTAATAAATCTAGATTTAGGGTGAAG 700
QY 161 LeuThrLeuGluGlyLeuGluAspAspAspArgValSerProThrValLeuHis 180
DB 701 CTGACACTAGAAAGGCTGGAGGAAGATGACGATGATAGGGTATCTCCCACTGTACTCCAC 760
QY 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGlnPheLysCysArgHis 200
DB 761 AAAATGTCCAATAGCTTGGAGATATCTTTAATAAGCGACAATGAGTTCAAGTGCAGGCAT 820
QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
DB 821 TCACAGCCGGAGTGTGGTTATGGCTTGCAGCCGTGATCGTTGGACAGAGTACAGCATACAG 880
QY 221 ThrMetGluProAspAsnLeuGluIlePheAspPheGluGluAspLeuSerGlu 240
DB 881 ACGATGGAACCAAGATTAACCTGGAACTAATCTTTTGTATTTTTCGAAGAGATCTCAGTGAG 940
QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260

DB 941 CACGTAGTTCCAGGTGATGCCCTTCTCTGGACATGTGGGTACAGCTTGTCTCTTATCATCC 1000
QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
DB 1001 ACCATTGCTGAGAGTGGAAAGAGTGTGGAAATTTCTTACTCTTCCATCATGACAGAAAT 1060
QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleLysProLeuProGly 300
DB 1061 TCCGGNAACAATAGCAAAAGTGAGAGTTGACTATATATATTATTAAGCCATTACCAAGGA 1120
QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
DB 1121 TACAGTTGTGACATGAATCTTTCATTTTCCAGTATTGGAAAGCCACAGAATACCATTCGAT 1180
QY 321 ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
DB 1181 GTTGGCCATCAGGTGCAGAAACTCTACAACTCCCGAGTGGCTAAAGTTCAAGAA 1240
QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
DB 1241 AATACTATTGCTCTTTAAGAAATGCTGCTAGTCAATGTCAGCCCTTTGTAGAATTGAC 1300
QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
DB 1301 GTACACCTTTCAAAGGACTTTGTGCCGTGGTATATCATGATCTTACCTGTGTTGACT 1360
QY 381 MetLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
DB 1361 ATGAAAAAGAAATTTGATGCTGCTCCAGTTGAATTTTGAATTTCCAGTAAAAAGAAATTA 1420
QY 401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
DB 1421 ACATTTACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACGTGAAATCTAAGAT 1480
QY 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
DB 1481 CGGAAAGAAATCTGGTTCAGAGGAAATTCCTTTTCAGAAATTCAGCCATTCTTCT 1540
QY 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
DB 1541 CTTAAGATGGTTTTAGAGTCTTTGCCAGAAAGATGAGGGTTTAAATTTGAATAAAAATGG 1600
QY 461 IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
DB 1601 ATCTGCCAGCAAAAGGATGGAATGTGGATGGTAACTTATCAACATATTTTGACATGAT 1660
QY 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal 500
DB 1661 CTGTTTTTGGATATATTTTAAAAAATCTTTTAAAGAAATCTGGGAAGAGAGAAATAGTG 1720
QY 501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro 520
DB 1721 TTTTCTTCATTTGATGCAGATATTTGCACATGTTTGGCAAAATCTCGGCAAGAGCAAAATATCCG 1780
QY 521 IleLeuPheLeuThrGlnGlySerGluIleTyrProGluLeuMetAspLeuArgSer 540
DB 1781 ATACTATTTTAACTCAAGGAAAAATCTGAGATTTATCTCTGAACCTCATGGACCTCAGATCT 1840
QY 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
DB 1841 CGGCAACCCCATTCATTCAGTATGAGCTTTGCAAGAGGCAAAAGCTAAAGGACTA 1900
QY 561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
DB 1901 CATACTGAAGACTTGCTCAGAAACCCATCTCTATATTCAAGAGGCAAAAGCTAAAGGACTA 1960
QY 581 VallLePheCysTrpGlyAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
DB 1961 GTCATATTCTCTGGGTGATGATACCAATGATCTCTGAAAAACAGAAAGAAATTTGAAGAA 2020
QY 601 LeuGlyValAlaAsnGlyIleLeuTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
DB 2021 CTTGAGTTAATGGTCTAAATTTATGATAGGATATATGATTTGGATGCCCTGAAACACCAAT 2080

QY 621 IlePheGlnValGluGlnLeuArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
 |||||
 Db 2081 ATATTCCAGTGGAGCAATTGGAAACGCTGAAGCAGCAATTGGCAGAGCTTAAGAGCTGT 2140
 |||||
 QY 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
 |||||
 Db 2141 TTGTGTCCCACTGTTAGCGCTTTGTCTCACTTTGTGTGGGAGTCTGATATCCAT 2200
 |||||
 QY 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
 |||||
 Db 2201 GTGGATGCCAACGGCATTGATACGTTGGAGATGCT 2236
 |||||

RESULT 5

CQ491863 3499 bp DNA linear PAT 30-JAN-2004
 LOCUS Sequence 23730 from Patent WO0160860.
 ACCESSION CQ491863
 VERSION CQ491863.1 GI:41457482
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
 TITLE Genes differentially expressed in human prostate cancer and their use
 JOURNAL Patent: WO 0160860-A 23730 23-AUG-2001;
 Millennium Predictive Medicine, Inc. (US)

FEATURES

source
 1..3499
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 9,86e-298 Length: 3499
 Score: 3537.00 Matches: 672
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-047-855-3 (1-672) x CQ491863 (1-3499)

QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
 |||||
 Db 221 ATGACACCTTCTCAGGTTCCTTTGAAATAAGAGGAACCTTTTACCAGGAGAGTCTTT 280
 |||||
 QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
 |||||
 Db 281 GCGATATGTGGAGCTGTGATGCTTTGGGAACTGGAACTCCCAAAATGCTGTGGCTCTT 340
 |||||
 QY 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
 |||||
 Db 341 CTTCCAGAGAATGACACAGGTGAAGCATGCTATGGAAGCAACCATTTGACTCAGTAGA 400
 |||||
 QY 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
 |||||
 Db 401 GGAGTATCAGTTTCAGTATCGCTACTTCAAAGGGTACTTTTAAAGCAACCAAGACTATCGGT 460
 |||||
 QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
 |||||
 Db 461 GGTCCATGTCAAGTAGATGTTCAAGTGGGAGACTCATCTACACCAAGCATCAATPACC 520
 |||||
 QY 101 ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
 |||||
 Db 521 CCTTTAGAAAGCGAATTATTATGACGATGGACAAATTGGAAATCCACAATGGTGTGAA 580
 |||||
 QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
 |||||

Db 581 ACTCTGGATTTCTGGATGGCTGACATGTCAGACTGAATAAGATTACGTTTGCATTATTCT 640
 |||||
 QY 141 GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys 160
 |||||
 Db 641 GAAAAACCTCTCTGTGTCAATAACCAAGAAAAAATAAAAAAATCTAGATTAGGGTGAAG 700
 |||||
 QY 161 LeuThrLeuGluGlyLeuGluLeuAspAspAspArgValSerProThrValLeuHis 180
 |||||
 Db 701 CTGACACTAGAGGGCTGGAGGAAGATGACGATAGTAGGGTATCTCCCACTGTACTCCAC 760
 |||||
 QY 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGlnPheLysCysArgHis 200
 |||||
 Db 761 AAAATGTCCAATAGCTTGGAGATATCTTAATAAGCACAATGAGTTCAGTGCAGGCAT 820
 |||||
 QY 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
 |||||
 Db 821 TCACAGCGGAGTGTGTTATGGCTTGACGCTGATCGTTGGACAGAGTACAGCATACAG 880
 |||||
 QY 221 ThrMetGluProAspAsnLeuGluIlePheAspPheGluGluAspLeuSerGlu 240
 |||||
 Db 881 ACGATGGAAACCAAGATAACCTGGAACTAATCTTTGATTTTTCGAAGAAGATCTCAGTGAG 940
 |||||
 QY 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
 |||||
 Db 941 CAGTAGTTACAGGTGATGCCCTTCTGGACATGGGGACAGCTTGCTCTTATCATCC 1000
 |||||
 QY 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
 |||||
 Db 1001 ACCATTGCTGAGAGTGGAAAGAGTGTGGAATTTCTTACTCTTCCCATCATGACAGAAAT 1060
 |||||
 QY 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGly 300
 |||||
 Db 1061 TCCCGGAAAAACAATAGGCAAGTGAAGTGTATATATATTAATTAAGGCATTTACCAGGA 1120
 |||||
 QY 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
 |||||
 Db 1121 TACAGTTGGACATGAATCTTCATTTTCCAAGTATTGGAAGCAAGAATACCATTTGGAT 1180
 |||||
 QY 321 ValGlyHisArgGlyValaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
 |||||
 Db 1181 GTTGGCCATCGAGGTGCAGGAAACTCTACAACAACCTGCCAGCTGGCTAAAGTTCAAGAA 1240
 |||||
 QY 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
 |||||
 Db 1241 AATACTATTGCTTCTTTAAGAAATGCTGCTAGTCTAGTGGTGCAGCCCTTGTAGAAATTGAC 1300
 |||||
 QY 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
 |||||
 Db 1301 GTACACCTTTCAAAGACTTTGTGCCCGGGTATATCATGATCTTACCTGTGTTGTGACT 1360
 |||||
 QY 381 MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu 400
 |||||
 Db 1361 ATGAAAAAGAAATTTGATGCTCATCCAGTTGAATTTATTGAAATTCAGTAAAGAAATTA 1420
 |||||
 QY 401 ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp 420
 |||||
 Db 1421 ACAATTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTGAAATCTAAGGAT 1480
 |||||
 QY 421 ArgLysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSer 440
 |||||
 Db 1481 CGGAAGAAATCTGTGGTTCAGGAGGAAAAATTCCTTTTCAGAAAAATCAGGCAATTTCTCT 1540
 |||||
 QY 441 LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp 460
 |||||
 Db 1541 CTTAAGATGTTTTAGAGTCTTTGGCAGGAAGATGAGGGTTTAAACATTGAATAAATG 1600
 |||||
 QY 461 IleCysGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn 480
 |||||
 Db 1601 ATCTGCACAAAGGAGTGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1660
 |||||
 QY 481 LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal 500
 |||||
 Db 1661 CTGTTTTTGGATATAATTTTAAAAAATCTGTTTTAGAAAAATTTCTGGGAAGGAGAAATAGTG 1720
 |||||

```
Qy 501 PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysThrPro 520
Db 1721 TTTTCTCATTTGATGCGAGATATTGCAATGGTTTCGGCAAAACGACAAATATCCG 1780
Qy 521 IleLeuPheLeuThrGlnGlnGlySerGluIleTyrProGluLeuMetAspLeuArgSer 540
Db 1781 ATACTATTTTAACTCAAGGAAATCTCGAGATTATCTCGAATCATCGGACCTCAGATCT 1840
Qy 541 ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnVal 560
Db 1841 CGGCAACCCCATTTGCAATGAGCTTTGCACAGATTGAAATCTACTGGGATAAATGTA 1900
Qy 561 HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu 580
Db 1901 CATACTGAAGACTTGCTCAGAAACCATCTCTATATTCAAGAGGGCAAAAGCTTAAGGGACTA 1960
Qy 581 ValIlePheCysTyrGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu 600
Db 1961 GTCATATTCTGCTGGGGTGATGATACCAATGATCTCTGAAACAGAGGAAATTTGAGGAA 2020
Qy 601 LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn 620
Db 2021 CTTCGAGTTAATGGTCTAATTATGATAGGATATATGATTGGATGCCTGAACACCAAT 2080
Qy 621 IlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCys 640
Db 2081 ATATTCCAAAGTGGAGCAATTGGAACGCTGAGCAGAGGAATTCGCCAGAGCTTAAGAGCTGT 2140
Qy 641 LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis 660
Db 2141 TTGTGTCCACCTGTAGCGCTTTGTTCCTCATCTTTGTGTGGGGAGTCTGATATCCAT 2200
Qy 661 ValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2201 GTGGATGCCAACGGCATTTGATAACGTGGAGATGCT 2236

RESULT 6
CQ497765 3499 bp DNA linear PAT 30-JAN-2004
LOCUS Sequence 29632 from Patent WO0160860.
DEFINITION CQ497765
ACCESSION CQ497765
VERSION CQ497765.1 GI:41463401
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 29632 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1..3499
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 9,86e-298 Length: 3499
Score: 2537.00 Matches: 672
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-047-855-3 (1-672) x CQ497765 (1-3499)
Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 1301 GTACACCTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTGTTGACT 1360
```

```
Db 221 ATGACACCTTCTCAGTTGCTTTGAAATAGAGGAACTCTTTTACCAGGAGAAAGTTTTT 280
Qy 21 AlalileCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 281 GCGATATGTGGAAGCTGTGATGCTTTGGAAACTGGAACTCTCAAAATGCTGTGGCTCTT 340
Qy 41 LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg 60
Db 341 CTTCCAGAGATGACACAGGTGAAAGCATGCTATGGAAAGCAACCATTTGTACTCAGTAGA 400
Qy 61 GlyValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
Db 401 GGAGTATCAGTTTCAGTATCGCTACTTCAAGGGTACTTTTAGAACCAACAGCATATCGGT 460
Qy 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
Db 461 GGTCCATGTCCTCAAGTGTAGATTTCACAGTGGGAGACTCATCTACACCAACAGATCAATACC 520
Qy 101 ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
Db 521 CTTTAGAAACGGAATTTATTATTGACATGAGCAATTTGGAAATCCACATGGTGTGAA 580
Qy 121 ThrLeuAspSerGlyTyrLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
Db 581 ACTCTGGATTCTGGATGGCTGACATGTCAGACTGAAATTAAGATTACGTTTGCATTATCT 640
Qy 141 GluLysProProValSerIleThrLysLysLysLysLysLysLysSerArgPheArgValLys 160
Db 641 GAAAAACCTCTGTGTCAATTAACCAAGAAAAAATTAATAAAATCTAGATTTAGGGTGAAG 700
Qy 161 LeuThrLeuGluGlyLeuGluAspAspAspArgValSerProThrValLeuHis 180
Db 701 CTGACACTAGAAAGCGCTGGAGGAAGATGACATGATAGGGTATCTCCACACTGTACTCCAC 760
Qy 181 LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 200
Db 761 AAATGTCCATAGCTTGGAGATATCTTATTAAGGACATAGATTCAAGTGCAGGCAT 820
Qy 201 SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGln 220
Db 821 TCACAGCGGAGTGTGGTTATGGCTTGAGCTGATCGTTGGACAGAGTACAGCATACAG 880
Qy 221 ThrMetGluProAspAsnLeuLeuIlePheAspPhePheGluGluAspLeuSerGlu 240
Db 881 ACAGTGGAAACAGATAACCTGGAATAATCTTTGATTTTTCGAAGAAGATCTCAGTGAG 940
Qy 241 HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer 260
Db 941 CACGTAGTTCAAGGTGATGCCCTTCCTGACATATGGGTACAGCTTGTCTCTATCATCC 1000
Qy 261 ThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsn 280
Db 1001 ACCATTGCTGAGAGTGGAAAGAGTGTGGAATCTTACTCTTCCCATCATGAGCAGAAAT 1060
Qy 281 SerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGly 300
Db 1061 TCCCGGAAAAACAATAGGCAAAAGTGAGCTTCACTATATAATTTATTAAGCCATTACAGGA 1120
Qy 301 TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp 320
Db 1121 TACAGTTGTGACATGAAATCTTCAATTTCCAAAGTATTTGAAGCCCAAGAAATACATTGGAT 1180
Qy 321 ValGlyHisArgGlyValAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu 340
Db 1181 GTTGGCCATCAGGTGCGAGGAACTCTACAACTGCCCAGCTGGCTAAAGTTCAAGAA 1240
Qy 341 AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp 360
Db 1241 AATACTATTGCTTCTTTAAGAAATGCTGCTAGTCACTGCTGAGCCCTTTGTAGAAATTGAC 1300
Qy 361 ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr 380
Db 1301 GTACACCTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTGTTGACT 1360
```

QY	381	MetLysGlyLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu	400
DB	1361	ATCAAAAGAAATTTGATGTCATCCAGTTGAATATTATTTGAATTCAGTAAAGAAATTA	1420
QY	401	ThrPheAspGlnLeuGlnLeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp	420
DB	1421	ACATTTGACCACTCCAGTTGTTAAAGCTACATCATGTGACTGCATGAATCTAAGGAT	1480
QY	421	ArgLysGluSerValValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSer	440
DB	1481	CGAAAGAAATCTGTGTTTCAGAGGAAATTCCTTTTCAGAAATCAGCAATTCCTTCT	1540
QY	441	LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluLeuLysTrp	460
DB	1541	CTTAAGATGTTTGTAGATCTTTGCCAGAAGATGTAGGGTTTAACATCGAAATAAATCG	1600
QY	461	IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTrpPheAspMetAsn	480
DB	1601	ATTCGCCAGCAAGGATGGAATGTGGATGGTAATCTTATCAACATATTTTGACATGAAT	1660
QY	481	LeuPheLeuAspIleLeuLysThrValLeuGluAsnSerGlyLysArgIleVal	500
DB	1661	CTGTTTTCGATATATTTTAAACTGTTTGAATAATCTCGGAGAGGAGATAGTG	1720
QY	501	PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTrpPro	520
DB	1721	TTTTTCTTCAATGATGATATTTTGCACAAATGTTCCGCAAAAGCAGCAAAATATCG	1780
QY	521	IleLeuPheLeuThrGlnGlySerGluIleTrpProGluLeuMetAspLeuArgSer	540
DB	1781	ATACTATTTTAACTCAAGGAAATCTGAGATTTATCTCGAACTCATGGACCTCAGATCT	1840
QY	541	ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal	560
DB	1841	CGGACACCCCATTCATGATGCTTTGGACATTTGAAATCTACTGGGATTAATGTA	1900
QY	561	HisThrGluAspLeuLeuArgAsnProSerTrpIleGlnGluAlaLysAlaLysGlyLeu	580
DB	1901	CATACTGAAGACTTGCTCAGAAACCCATCCTATATTCAAGAGGCAAAAGCTAAGGACTA	1960
QY	581	ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu	600
DB	1961	GTCAATATCTGCTGGGGTGTATGATACCAATGATCTCTGAAACACAGAGGAAATGGAAGAA	2020
QY	601	LeuGlyValAsnGlyLeuIleTrpAspArgIleTrpAspTrpMetProGluGlnProAsn	620
DB	2021	CTTGGATTTAATGGTCTAATTTATGATAGATATATGATGGATGCTGGAACACCAAT	2080
QY	621	IlePheGlnValGluGlnLeuGluArgLysGlnGluLeuProGluLeuLysSerCys	640
DB	2081	ATATTCCAAGTGGAGCAATTGGAACGCTGAAGCAGCAATTCGACAGCTTAAGAGCTGT	2140
QY	641	LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHis	660
DB	2141	TTGTGTCCCACTGTTAGCCGCTTTGTTCCCTCATCTTTGTGGGGAGTCTGATATCCAT	2200
QY	661	ValAspAlaAsnGlyIleAspAsnValGluAsnAla	672
DB	2201	GTGGATGCCAACGGCAATTGATACGTGGAGATGCT	2236
RESULT 7			
AB037855			
LOCUS			
AB037855 Homo sapiens mRNA for KIAA1434 protein, partial cds.			
DEFINITION			
AB037855			
ACCESSION			
AB037855.1 GI:7243265			
VERSION			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (sites)			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M., and Ohara,O.			
Prediction of the coding sequences of unidentified human genes.			
XVI. The complete sequences of 150 new cDNA clones from brain which			
code for large proteins in vitro			
DNA Res. 7 (1), 65-73 (2000)			
20181126			
10718198			
2 (bases 1 to 5443)			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
Ohara,O., Nagase,T. and Kikuno,R.			
Direct Submission			
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,			
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba			
292-0812, Japan [E-mail:cdmainfo@kazusa.or.jp]			
URL:ftp://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,			
Fax:+81-438-52-3914			
FEATURES			
Location/Qualifiers			
1..5443			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="hh05877"			
/tissue_type="brain"			
/clone_lib="pBluescriptII SK plus"			
1..5443			
/gene="KIAA1434"			
<190..2223			
/gene="KIAA1434"			
/note="Start codon is not identified."			
/codon_start=1			
/product="KIAA1434 protein"			
/protein_id="BA092672.1"			
/db_xref="GI:7243266"			
/translation="NKSIRMTSQVAFEIRGLLPGEVFAICGSCDALGNWPNQNAVA			
LLPNDTGESMLWKATIVLSRGVSQVYRFGYFLEPKTIGGPOVIVHKWTHLQPR			
STTPLESIIIDGOGIHNGVETLDSGLTCTQTEIRLRHLYSEKPPVSIITKKLKS			
RFRVLTLEGLEEDDDRSPTVLHMSNSLEISLIDNEFKRHSQPGCGVGLQPR			
WTEYSIQWENLESLIDFFEDSESHVQGDALPGHVTACLISSTIAESGKSAGI			
LTPILMSNRKTIQKRVYIIIIKPLPGYSCKMSKSFYKWKPRILPLDVHGRGAGNS			
TTTAQLAKQENTIASLRNAAASHGAFFEDVHLKDFVYVYHDTCCLTWKKKFD			
DVEUFEIPKELTDLQLLKTHTVLTAKSKDRKESVVOEENSFSENPFPRLKML			
BSLPDVGFIKWKICQORDGMWGNLSTYFDMNLFLDIILKTVLENGKRIVFSS			
EDADICTWVRQKQKYPILFLTGKSEIYPELMDLRSRTPIAMSPAQENILGINVH			
TEDLUNPSYIOEAKAGLVIFCWGDDTNDPNRRKLKELGVNGLIYDRIYDWMRPQP			
NIFQVEQLERLKLPELKLCTPTVSRFVPSLSCGESDIHVNDANGIDNVENA"			
ORIGIN			
Alignment Scores:			
Pred. No.: 1.79e-297 Length: 5443			
Score: 3537.00 Matches: 672			
Percent Similarity: 100.00% Conservative: 0			
Best Local Similarity: 100.00% Mismatches: 0			
Query Match: 100.00% Indels: 0			
DB: 9 Gaps: 0			
US-10-047-855-3 (1-672) x AB037855 (1-5443)			
QY	1	MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe	20
DB	205	ATGACACCTTCTCAGGTTGCCTTTGAAATAAGAGGACTCTTTTACCAGAGAAGTTT	264
QY	21	AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu	40
DB	265	GGCATATGTGGAAGCTGTGATGCTTTGGGAAACCTGGAATCCTCAAAATGCTGGCTCT	324
QY	41	LeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleValLeuSerArg	60
DB	325	CTTCAGAGAATGACACAGGTGAAGCATGCTATGGAAGCAACCACTTGATCTCAGTAGA	384
QY	61	GlyValSerValGlnTrpArgTrpPheLysGlyTrpPheLeuGluProLysThrIleGly	80
DB	385	GGAGTATCAGTTCAGTATCGCTACTTCAAGGGGTACTTTTGTAGAACCAAGACTACGGT	444
QY	81	GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr	100

Db	445	GGTCCATGTCAGTATAGTTCACAGTGGGAGCTCATCTCAACACCAATCAATAACC	504	Db	1525	CTTAAAGATGGTTTTAGAGTCTTTGCCAGAGATGTAGGGTTTAAACATTGAAATAAAATGG	1584
Qy	101	ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu	120	Qy	461	IleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsn	480
Db	505	CTTTTAGAAACGCAAAATTATTATTGACGATGGACAAATTGGAAATCCCAATATGGTGTGAA	564	Db	1585	ATCTGCCACCAAGGATGGATGGTAACTTATCAACATATTTTGCACATGAAT	1644
Qy	121	ThrLeuAspSerGlyTyrLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer	140	Qy	481	LeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleVal	500
Db	565	ACTCTGATTTCTGATGGCTGACATGTGACACTGAAATAAGATTACCTTTGGCATTTCT	624	Db	1645	CTGTTTTTGGATATATATTTTAAAAACGTGTTTAGAAAAATTTCTGGGAAGAGGAGATAGTG	1704
Qy	141	GluLysProProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLys	160	Qy	501	PheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrPro	520
Db	625	GAATAAACTCTCTGTCTCAATAACCAAGAAAAATTAATAAATCTAGATTAGGGTGAAG	684	Db	1705	TTTCTCTTCAATTTGATCAGATATTTGCACAATGGTTGGCAAAAGCAGAAACAATATCCG	1764
Qy	161	LeuThrLeuGluGlyLeuGluGluAspAspArgValSerProThrValLeuHis	180	Qy	521	IleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSer	540
Db	685	CTGACACTAGAAAGCCCTGGAGGAAGATGACGATGATAGGGTATCTCCCACTGTACTCCAC	744	Db	1765	ATACTATTTTAACTCAGGAAATCTGAGATTTATCTCTGAACTCATGGACCTCAGATCT	1824
Qy	181	LysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis	200	Qy	541	ArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuGlyIleAsnVal	560
Db	745	AAATGTCCAATAGCTTGGATATCTCTTAATAAGCGACAATGATGATCAAGTGCAGGAT	804	Db	1825	CGGACAAACCCCATTTGCAATGAGCTTTGCACAGTTTGAATAATCTACTGGGGATAAATGTA	1884
Qy	201	SerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTyrThrGluTyrSerIleGln	220	Qy	561	HisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeu	580
Db	805	TCACAGCCGAGTGGTATGGCTTGCAGCTGATCGTTGGACAGAGTACAGATACAG	864	Db	1885	CATACTGAAGACTTGTCTGAAACCCCATCTATATTCAAGAGGCAAAAGCTAAGGACTA	1944
Qy	221	ThrMetGluProAspAsnLeuGluIlePheAspPheGluGluAspLeuSerGlu	240	Qy	581	ValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGlu	600
Db	865	ACGATGGAAACAGATAACTGGAACTATCTTGATTTTTCGAAGAAGATCTCAGTGAG	924	Db	1945	GTCAATATTTCTGCTGGGTGATGATCAACATGATCTCTGAAACAGAGAAATTTGAAGAA	2004
Qy	241	HisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSer	260	Qy	601	LeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGlnProAsn	620
Db	925	CAGTAGTTCCAGGTGATGCCCTTCTCGACATGTGGGTACAGCTTGTCTCTATCATCC	984	Db	2005	CTTGGAGTTAATGGTCTAATTTATGATAGGATATATGATGGATGCTGCAACCAACCAAT	2064
Qy	261	ThrIleAlaGluSerGlyLysSerIleGlyIleLeuThrLeuProIleMetSerArgAsn	280	Qy	621	IlePheGlnValGluGlnLeuGluArgLeuLysGlnLeuLeuProGluLeuLysSerCys	640
Db	985	ACCATTGCTGAGAGTGGAAAGAGTCTGGAATCTTACTCTTCCCATCATGAGCAGAAAT	1044	Db	2065	ATATTCCAAGTGGAGCAATTTGGAACGCTGAAGCAGAAATGGCAGAGCTTAAGAGCTGT	2124
Qy	281	SerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGly	300	Qy	641	LeuCysProThrValSerArgPheValProSerSerLeuCysGlyGlyGluSerAspIleHis	660
Db	1045	TCCCGGAAACAAATAGGCAAGTGAAGTGAAGTGAATATAATATTAAAGCCATTACAGGA	1104	Db	2125	TTGTGTCCTCCACTGTGATGCCGCTTTGTCCTCATCTTTGTGTGGGAGTCTGATATCCAT	2184
Qy	301	TyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAsp	320	Qy	661	ValAspAlaAsnGlyIleAspAsnValGluAsnAla	672
Db	1105	TACAGTTGTGACATGAATCTTCAATTTCCAAAGTATTGGAAGCCCAAGAAATACATTGGAT	1164	Db	2185	GTGATGCCAAGCGCATTTGATAACGTGAGAAATGCT	2220
Qy	321	ValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGlu	340	RESULT 8			
Db	1165	GTGGCCATCGAGGTGCAGAAACTCTACAACTGCCCAGCTGGCTAAAGTTCAAGAA	1224	AY233980	3229 bp	mrna	linear ROD 06-APR-2003
Qy	341	AsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAsp	360	LOCUS			
Db	1225	AAATCTATTGCTTCTTTAAGAATGCTGCTAGTCAATGGTGCAGCTTTGTAGAAATTTGAC	1284	DEFINITION	Rattus norvegicus hypothetical protein LK44	mrna, complete cds.	
Qy	361	ValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThr	380	ACCESSION	AY233980		
Db	1285	GTACACCTTTCAAAGGACTTTGTGCCCGTGGTATATCATGATCTTACCTGTTGTTGACT	1344	VERSION	AY233980.1	GI:29569150	
Qy	381	MetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLysGluLeu	400	KEYWORDS			
Db	1345	ATGAAAGAAATTTGATGCTGATCCAGTTGAATATTTTGAAATTCAGTAAAGAAATTA	1404	SOURCE	Rattus norvegicus (Norway rat)		
Qy	401	ThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAsp	420	ORGANISM	Rattus norvegicus		
Db	1405	ACATTTGACCACTCAGTTGTTAAAGCTCCTCATGTGATCGCATGAAATCTAAGGAT	1464	REFERENCE	1 (bases 1 to 3229)		
Qy	421	ArgLysGluSerValValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSer	440	AUTHORS	Li, J. Y., Boado, R. J. and Partridge, W. M.		
Db	1465	CGGAAAGAAATCTGCTGCTCAGGAGGAAATTCCTTTTCAGAAATCAGCAATTCCTCTCT	1524	TITLE	Blood-brain barrier genomics		
Qy	441	LeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrp	460	JOURNAL	J. Cereb. Blood Flow Metab. 21 (1), 61-68 (2001)		
				MEDLINE	21024389		
				PUBMED	11149669		
				REFERENCE	2 (bases 1 to 3229)		
				AUTHORS	Li, J. Y., Boado, R. J. and Partridge, W. M.		
				TITLE	Direct Submission		
				JOURNAL	Submitted (11-FEB-2003) Medicine, UCLA, 900 Veteran Ave., Warren		
				FEATURES	Hall 13-164, Los Angeles, CA 90024, USA		
				source	Location/Qualifiers		
					1..3229		
					/organism="Rattus norvegicus"		
					/mol_type="mrna"		

/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/issue_type="brain microvasculature"
108. .2126
/notes="enriched at the brain microvasculature; similar to
Homo sapiens hypothetical protein in GenBank Accession
Number XM_045585"
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein LK44"
/protein_id="AA084024.1"
/db_xref="GI:29569151"
/translation="MTPSOVTEIRGTLLPGEVFMCGNCDALGNWSPQNAVPLTSE
TESVWKAIVISRGMSVKYRFRGCFLEPKTIGGPCQVIVHKWETHLQPRSIITPLEN
EIIIDQGIHNGVETSLKGLTCQTEIRLHLHFSEKPPVSIITKKPKKSRFRVKLT
LEGLEDDDDDKASPTVLHKMSNLISLISDNFKRHSQPECGYGIQPRWTEYS
IOTMEPDNLELIFDPEEDLSHVQDVLPGHVGTACLLSSTIAESBSRSGILLTUPI
MSRSRKTIGKVRVDPIIKPLPGVSCSMQSFYSKYWKPRIPLDVCHRGAGNSTTAK
LAKVQNTIASLNAASHGAAVFEFVDHLSKOLVPVYHDLTCCLTKRKYEADPVEL
FEIPKELTFDQLQULKSHVLTAKTDQKQWAEENESFSENQPPFSKKRMVLESLE
NVGFNEIKWI COHRDGVNDGLSTYFDNNAFLDIILKTVLENSGKRIRVFSFSDADI
CTWROKQNKYPIILFTQKSDIYPELMDLRSTPIAMSFQENILGINAHTEDLL
RNPSYVQEAQKGLVIFCWGDDTNDPENRRKLKEFGVNGLIYDRIYDMWPEQNPFOV
EQLERLKLPELKNCLCTVSHFIPPSFCMESKIHVDANGIDNVENA"

ORIGIN

Alignment Scores:
Pred. No.: 3,936-273 Length: 3229
Score: 3254.00 Matches: 613
Percent Similarity: 95.99% Conservative: 34
Best Local Similarity: 90.95% Mismatches: 23
Query Match: 92.00% Indels: 4
DB: 10 Gaps: 3

US-10-047-855-3 (1-672) x AY233980 (1-3229)

QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
DB 108 ATGACACCTTCTCAGGTCACCTTTGAAATAAGAGGAACCTTTTACAGAGAGGCTCTTT 167
QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
DB 168 GCAATGTGGAAACATGTGATGCTTGGGAAACCTGGAGTCTCTCAAAATGCTGTCGCT--- 224
QY 41 LeuProGlnAsnAspThrGlyGluSerMetLeuTrpIleAlaThrIleValLeuSerArg 60
DB 225 CTATCTGAGTGGAGTACAGCGGAAAGT---GTATGGAAAGCAGTGTATGTTCTTAGTAGA 281
QY 61 GlyValSerValGlnTrpArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGly 80
DB 282 GGAATGTCGGTGAAGTACCGCTACTTTCAGAGGCTGCTTTTGAACCAAGACTATCGGT 341
QY 81 GlyProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThr 100
DB 342 GGTCCATGTCAAGTCATAGTTCACAGTGGAGACTCATCTACACCCAGCATCAATAACC 401
QY 101 ProLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGlu 120
DB 402 CCTTTAGAAACGAATCAATATTGACGATGGACAAATTTGGAATCCACATGGTGTTGAA 461
QY 121 ThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSer 140
DB 462 ACATCGATTCTGGATGGCTTACCTGTACAGCTGAAATAAGACTGCGTCTGCATTTTCT 521
QY 141 GluLysProValSerIleThrLysLysLysLysLysSerArgPheArgValLys 160
DB 522 GAGAAACCTCTGTTTCAATATACCAAGAAAAAGTTCAAAAAATCTAGATTAGGGTAAAG 581
QY 161 LeuThrLeuGluGlyLeuGlu-----AspAspAspAspArgValSerProThrVal 178
DB 582 CTTACACTAGAGGCTCTGAGGAGAGATGATGACGATGATGATGATGATGATGATGATGAT 641
QY 179 LeuHisLysMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCys 198

DB 642 CTTTCAAGATGTCCAAATAGCTGGAGATATCTTAAATAGTGACAATGAGTTCAGATGC 701
QY 199 ArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSer 218
DB 702 AGGCACTCACAGCCAGNATGTGGTATGGCTTACAGCTTGACGCTGGACAGGTACAGC 761
QY 219 IleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPhePheGluGluAspLeu 238
DB 762 ATACAGACAATGGAGCGGACAACCTTGAACCTCATCTTTGACTTTTTCGAGGAAGATCTC 821
QY 239 SerGluHisValValGlnIleAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeu 258
DB 822 AGTGAGCATGTAGTCAGGTCATGTTCTTCTGGACATGTGGGCACACGATGCCCTCGT 881
QY 259 SerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSer 278
DB 882 TCATCTACCATTTGCTGAGAGTGAGAGAGCGGTGGATCTTCTTCTTCCCATCATGAGC 941
QY 279 ArgAsnSerArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeu 298
DB 942 AGAAGTTCAGAAAAAATAAGCAAAAGTCAGAGTTGATTTTATCATCATCAAGCCATTA 1001
QY 299 ProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIlePro 318
DB 1002 CCAGGATATAGTTGTTCTATGAGTCTTCTTCTTCAAGTATTTGGAAAAACAAGATACCA 1061
QY 319 LeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysVal 338
DB 1062 CTGGATGTTGGACATCGTGTGTCAGGAACTCAACAACTGCGCAAGTGGCTAAAGTA 1121
QY 339 GlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGlu 358
DB 1122 CAGGAAAAATCTATTGCTTCTTTAAGAAATGCTGCCAGCATGTTGGGAA 1181
QY 359 PheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCys 378
DB 1182 TTTGATGTCCACTTTTCAAGGACTTAGTGCTGTAGTGTATCATCATCTCCTCTGT 1241
QY 379 LeuThrMetLysLysLysPheAspAlaAspProValGluLeuPheGluIleProValLys 398
DB 1242 TTAACATGAAAGGAATATATAGCTGATCCAGTTGAATGTTTGAATCCACGATTAAG 1301
QY 399 GluLeuThrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSer 418
DB 1302 GAATTAACATTCGACCAACTCCAGTTATTGAAAGCTTCTCATGTGATGCTCACTAAAAACC 1361
QY 419 LysAspArgLysGluSerValValGlnGluGluAsnSerPheSerGluAsnGlnProPhe 438
DB 1362 AAAGACCAGAAACAATGTATGGCTGAGGAGGAAAAATTCCTTTCTGAAAAACCAACCATTT 1421
QY 439 ProSerLeuLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIle 458
DB 1422 CTTTCTTTAAGATGTTTTAGATCATTCGACGAAATGTTAGGATTTAATATAGAAATA 1481
QY 459 LysTrpIleCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAsp 478
DB 1482 AAATGGATTTGCAACACAGGATGGAGTATGGAGCGCAACTTATCGACATATTTTGTAT 1541
QY 479 MetAsnLeuPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArg 498
DB 1542 ATGAATGCATTTTGGATATAATTTAAAAAATGTTTTAGAAAAATTCGGAAGAGGAGA 1601
QY 499 IleValPheSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLys 518
DB 1602 ATAGATTTTCTTCAATTTGATGACATCTGTACAAATGGTTCGGCAGAAAAACAAACAA 1661
QY 519 TyrProIleLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeu 538
DB 1662 TATCCATATATTTTGGACCAAGGAAGTCTGACATTTACCTGAACCTCATGGACCTC 1721
QY 539 ArgSerArgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIle 558

```

Db 1722 AGATCTCGACAAACACCCATTGCATGAGCTTTCACAGCTTGAATAATATTTGGGATA 1781
Qy 559 AsnValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysAlaLys 578
Db 1782 AATCCCATACTAGATCTCTTAGAAACCATCTATGTCCAAAGAGGCAAAAGATAAG 1841
Qy 579 GlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeu 598
Db 1842 GGATTGGTCATATCTCTGGGGTGATGATACCAATGATCTCGAAACACAGAAGAAACTG 1901
Qy 599 LysGluLeuGlyValAsnGlyLeuIleTyrAspArgIleTyrAspTrpMetProGluGln 618
Db 1902 AAGGAATTTGGAGTAAATGGTCTAATATATGATAGATATACGATTGGATGCCGTGAACAA 1961
Qy 619 ProAsnIlePheGlnValGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLys 638
Db 1962 CCAATATATTTCCAAAGTGGAGCACTGGAGCGCTGAAGCGAGAATTGCCAGAGCTTAAG 2021
Qy 639 SerCysLeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAsp 658
Db 2022 AACTGTTTGTGCCACTGTTAGCCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2081
Qy 659 IleHisValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 2082 ATCCATGTGGATGCTAACGGCATTGATAATGTGGAGAACGCT 2123

RESULT 9
AC144439 39694 bp DNA linear HTG 16-APR-2003
LOCUS Homo sapiens chromosome 7 clone XXFOS-87042G10, WORKING DRAFT
DEFINITION SEQUENCE, 3 unordered pieces.
ACCESSION AC144439
VERSION AC144439.1 GI:29893792
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 39694)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 39694)
Waterston,R.H.
Direct Submission
Submitted (16-APR-2003) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: H AA087042G10
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 39167 bases at least Q40
Consensus quality: 39319 bases at least Q30
Consensus quality: 39366 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

```

* as soon as it is available and the accession number will
* be preserved.
* 1 7028: contig of 7028 bp in length
* 7029 7128: gap of unknown length
* 7129 16822: contig of 9694 bp in length
* 16823 39694: gap of unknown length
* 16923 39694: contig of 22772 bp in length.

FEATURES
    source
        1..39694
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="7"
            /clone="XXFOS-87042G10"
    misc_feature
        1..7028
            /note="assembly_name:Contig4"
    misc_feature
        7129..16822
            /note="assembly_name:Contig5"
    misc_feature
        16923..39694
            /note="assembly_name:Contig6"

ORIGIN
Alignment Scores:
Pred. No.: 6.89e-249 Length: 39694
Score: 2992.50 Matches: 611
Percent Similarity: 78.50% Conservative: 17
Best Local Similarity: 76.38% Mismatches: 35
Query Match: 84.61% Indels: 139
DB: 2 Gaps: 6

US-10-047-855-3 (1-672) x AC144439 (1-39694)
Qy 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
Db 32454 ATGACACCTTCTCAGGTTACCTTTGAATAATAGAGAACTCTTTTACCA---GAAGTTTT 32510
Qy 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
Db 32511 GCATATGTGGAAGCTGTGATGCTTTGGAAACTGGAATCTCTCAAAATGCTGTGGCTCTT 32570
Qy 41 LeuProGluAsnAspThrGly-Glu----- 48
Db 32571 CTTCCAGAGATGAGACAGAGGTGGAAGAGCTTAAGATTGTATCAGAAATGAGGGCAATGCC 32630
Qy 49 -----SerMetLeuTrpIleAlaThrIleValLeuSerArgL 61
Db 32631 TAAACCTGCTTTAGCAGAGCCAGCATGTTATGAAAGCAACCATTTGACTCAGTAGAGG 32690
Qy 61 yValSerValGlnTyrArgTyrPheLysGlyTyrPheLeuGluProLysThrIleGlyL 81
Db 32691 AGTATCAGTTCAGCATTCCTACTTCAAGGGGTACTTTTAGAACCAAGACTATCGGTGG 32750
Qy 81 yProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThrPr 101
Db 32751 TCCATGTCAAGTGATAGTTTCAAGTGGGAGACTCATCTACACCAACCATCAATAACCCC 32810
Qy 101 oLeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGluTh 121
Db 32811 GTTAGAAAGTGAATTTGTTATTGACCATGGACAATTTGGAATCCACAGTGGTGTGAAC 32870
Qy 121 rLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerGl 141
Db 32871 TCTGGATTCTGCATGCTGGCATGTCAGATCAATAATAGATTACGATTGCATTATTCTGA 32930
Qy 141 uLysProProValSerIleThrLysLysLeuLysLysSerArgPheArgValLysLe 161
Db 32931 AAAACCTCTGTGTCAATAACCAAGAAAAATTTTAAAAATCTAGATTTAGGGTGAAGCT 32990
Qy 161 uThrLeuGluGlyLeuGluGluAspAspAspArgValSerProThrValLeuHisly 181
Db 32991 GACACTAGAGGGCCTGGAGGAAGATGAGGATGATAGGGTATCTCCACTGTACTTACAA 33050
Qy 181 sMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 201

```



```
REFERENCE
AUTHORS
TITLE
JOURNAL

4 (bases 1 to 149259)
Waterston,R.H.
Direct Submission
Submitted (28-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 149259)
Waterston,R.
Direct Submission
Submitted (03-APR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 149259)
Waterston,R.H.
Direct Submission
Submitted (13-APR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
7 (bases 1 to 149259)
Waterston,R.
Direct Submission
Submitted (20-APR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
8 (bases 1 to 149259)
Wilson,R.
Direct Submission
Submitted (08-OCT-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Apr 13, 2002 this sequence version replaced gi:19774659.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH1217F02
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/FTP/CHR7, send
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6
Location/Qualifiers
1. .149259
/organism="Homo sapiens"
/mol_type="genomic DNA"

FEATURES
source
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP11-1217F2"
/clone_lib="RPCI-11"
1. .61
/rpt_family="Alu"
repeat_region
466. .658
/rpt_family="L1"
repeat_region
678. .978
/rpt_family="Alu"
repeat_region
2063. .2155
/rpt_family="MER1_type"
repeat_region
2352. .2466
/rpt_family="MIR"
repeat_region
2592. .2938
/rpt_family="L1"
repeat_region
2939. .3225
/rpt_family="Alu"
repeat_region
3226. .3326
/rpt_family="L1"
repeat_region
3327. .4018
/rpt_family="L1"
repeat_region
4021. .4339
/rpt_family="Alu"
repeat_region
4340. .4602
/rpt_family="ERV1"
repeat_region
4621. .5406
/rpt_family="ERV1"
repeat_region
5441. .5500
/rpt_family="AT_Rich"
repeat_region
5541. .5597
/rpt_family="MaLR"
repeat_region
5598. .5886
/rpt_family="Alu"
repeat_region
5887. .6209
/rpt_family="MaLR"
repeat_region
6251. .6324
/rpt_family="MER1_type"
repeat_region
6325. .6621
/rpt_family="Alu"
repeat_region
6622. .6810
/rpt_family="MER1_type"
repeat_region
6825. .7400
/rpt_family="ERV1"
repeat_region
7406. .7702
/rpt_family="Alu"
repeat_region
7719. .7786
/rpt_family="ERV1"
repeat_region
7792. .8139
/rpt_family="ERV1"
repeat_region
8140. .8443
/rpt_family="Alu"
repeat_region
8444. .8832
/rpt_family="ERV1"
repeat_region
8842. .9157
/rpt_family="Alu"
repeat_region
9189. .9677
/rpt_family="ERV1"
repeat_region
9678. .9764
/rpt_family="Alu"
repeat_region
9765. .9907
/rpt_family="ERV1"
repeat_region
9924. .10204
/rpt_family="Alu"
repeat_region
10240. .10284
/rpt_family="ERV1"
repeat_region
10363. .12161
/rpt_family="BSR/Beta"
repeat_region
12247. .13389
/rpt_family="BSR/Beta"
repeat_region
13417. .14673
/rpt_family="BSR/Beta"
```

```
repeat_region 14714..15351
/rpt_family="BSR/Beta"
repeat_region 15443..15502
/rpt_family="ERV1"
repeat_region 15503..15792
/rpt_family="Alu"
repeat_region 15793..16304
/rpt_family="ERV1"
repeat_region 16322..16656
/rpt_family="ERV1"
repeat_region 16658..17243
/rpt_family="ERV1"
repeat_region 17244..17312
/rpt_family="(TTTCC)n"
repeat_region 17314..17478
/rpt_family="L1"
repeat_region 17508..17818
/rpt_family="Alu"
repeat_region 18100..18160
/rpt_family="L1"
repeat_region 1849..19151
/rpt_family="Alu"
repeat_region 19152..19535
/rpt_family="L1"
repeat_region 20458..20489
```

Alignment Scores:

```
Pred. No.: 4.1e-248 Length: 149259
Score: 2992.50 Matches: 611
Percent Similarity: 78.50% Conservative: 17
Best Local Similarity: 76.38% Mismatches: 35
Query Match: 84.61% Indels: 139
DB: 9 Gaps: 6
```

US-10-047-855-3 (1-672) x AC099654 (1-149259)

```
QY 1 MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
DB 23087 ATGACACCTTCTCAGGTATACCTTTGAATAAGAGGAACCTCTTTTACCA---GAAGTTTTT 23031
QY 21 AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
DB 23030 GCAATATGTGGAAGCTGTGATGCTTTGGGAACTGGAACTCCTCAAAATGCTGTGGCTCT 22971
QY 41 LeuProGluAsnAspThrGly-Glu----- 48
DB 22970 CTTCCAGAGNATGAGACGAGTGGGAAGACTAAGATTGTATCAGATGAGGGCAATGCC 22911
QY 49 -----SerMetLeuTrpLysAlaThrIleValLeuSerArgG 61
DB 22910 TAAACCTGCTCTAGCAGAGCCAGCATGTTATGGAAAGCAACCATTTGTACTAGTAGAG 22851
QY 61 yValSerValGlnTrpArgTyPheLysGlyTyPheLeuGluProLysThrIleGlyG 81
DB 22850 AGTATCAGTTCAGCATGTGCTCTTCAAGAGGTACTTTTTAGAACCAAGACTATCGTGG 22791
QY 81 yProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThrPr 101
DB 22790 TCCATGTCAGTGATAGTTTCAAGTGGGAGACTCATCTACAACACGATCAATAACCC 22731
QY 101 leuGluSerGluIleIleleaspaspGlyGlnPheGlyIleHisAsnGlyValGluTh 121
DB 22730 GTTAGAAAGTGAATTTGTTATGACCATGGACAAATTTGGAATTCACAGTGGTGTGAAAC 22671
QY 121 rLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerG 141
DB 22670 TCTGGATTCTGCATGCTGGCATGTCAGACTGGAATTAAGATTACGATTGCAATTCTCGA 22611
QY 141 ubysProProValSerIleThrLysLysLysLysSerArgPheArgValLysLe 161
```

```
DB 22610 AAAACCTCTGTGTCAATAACCAAGAAAAATTTTAAAAATCTAGATTTAGGGTAGAGCT 22551
QY 161 uThrLeuGluGlyLeuGluGluAspAspArgValSerProThrValLeuHisLys 181
DB 22550 GACACTAGAGGGCTGGAGGAAGATGAGGATATCTCCCACTGCTACTTCACAA 22491
QY 181 sMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHis 201
DB 22490 AATGTCCAATAGCATGGAGATATCTTAATATATGACAATAGCTTCAGTCGAGGATTC 22431
QY 201 rGlnProGluCysGlyTyPheGluGlnProAspArgTrpThrGluTySerIleGlnTh 221
DB 22430 ACAGCTGGAGTGTGGTTATGGCTTACAGTCTGATCTGTGGACAGATACAGCATACAG 22371
QY 221 rMetGluProAspAsnLeuGluLeuIlePheAspPheGluGluAspLeuSerGluHi 241
DB 22370 AATGGGACCAGATTAACCTGGAACTAATCTTGAATTTTGAAGAAGATCTCAGTGAGCA 22311
QY 241 sValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerTh 261
DB 22310 CATAGTTCAAGGTGATGCCCTCCCTGGATGTGGGTACAGCTTGTCTCTATATCATCCAC 22251
QY 261 rIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsnSe 281
DB 22250 CATGTCTGAGAGTGGAAAGAGTGTGGAATTTCTTACTTCTCCCATCATGAGCAGAATTC 22191
QY 281 rArgLysThrIleGlyLysValArgValAspTrpIleIleIleLysProLeuProGlyTy 301
DB 22190 CAGGAAACAAATAGGCAAGTGAAGTGAATATATATATTTAAGCAATGCGCAGATA 22131
QY 301 rSerCysAspMetLysSerSerPheSerLysTyTrpLysProArgIleProLeuAspVa 321
DB 22130 CAGTTGTGACATGAATCTTCAATTTTCCAG-TATTGGAAAGCCAAAGAAATACCATGATGT 22072
QY 321 lGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAs 341
DB 22071 TGGCCATCAGAGTGTGGGAAACGCT---ACAACTGCCAGCTGGCTGCTAAA-----GAAAA 22021
QY 341 nThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAspVa 361
DB 22020 TACTATTGCTTCTTTTAAGAAATGCTGTAGTACGCTGAGGCTTTGTAGATTGATGT 21961
QY 361 lHisLeuSerLysAspPheValProValValTyHisAspLeuThrCysCysLeuThrMe 381
DB 21960 ACACCTTTCAAAGGACTTTGTGCCATGCTATATCATGATCTTACCTGTTGTTGACTAT 21901
QY 381 tLysLysLysPheAspAlaAspProValGluLeuPheGluIleProVal-LysGluLeuT 401
DB 21900 GAAAAAGAAATTTGATGCTGATCCAGTTGAATTTATTTGAATTTCCAGTAAAAAATTA 21841
QY 401 hrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAspA 421
DB 21840 CATTTGACCATCTCCAGTTGTTAAAGCTCCTCATGTGACTGCACGTGAATCTTAAGATT 21781
QY 421 rGlySerGluSerValValGlnGluGluAsnSerPheSerGluAsnGlnProPheProSerL 441
DB 21780 GGAAGAATCTGTGGTCCAGGAGGAAATTCCTTTTCAGAAAAATCAGGCAATTTCTCTC 21721
QY 441 euLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpI 461
DB 21720 GTAAGATGTTTTTAGAGTGTTTGCCAGGAAGATGTAGGGTTTAACATTTGAATAAATGA 21661
QY 461 leCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyPheAspMetAsnL 481
DB 21660 TCTGCCAGCAAAAGGATGGAATGTGGGATGGTAACCTTATCAACATATTTTGGCATGAATC 21601
QY 481 euPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValP 501
DB 21600 TGTTTTGGATATACTTTTAAAACTGTTTAAAGAAATTTTGGAAAGAG-AGAAATAGTGT 21542
QY 501 heSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyProI 521
```


Db	21541	TTTCTTCAATTGATGCAGATATTTCACAATGGTTTGGCAAGGCGAGCAAAATATCCCA	21482
Qy	521	leLeuPheLeuThrGlnGlySerGluIleTyProGluLeuMetAspLeuArgSerA	541
Db	21481		
Db	21481	TATTATTTTAACTGAAGAAATCTGAGATTATCTCTGAACTCATGCACCTCAGATCTC	21422
Qy	541	rgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeu-	556
Db	21421	CGACAACCCCATTTGCAATGAGCTTTGCACAGTTTGAAATCTACTAGGGATAGGCTGGG	21362
Qy	556	-----	556
Db	21361	TGCGGTGCTCGCGCTGTAAACCCAGCACATTTGGGAGCGCGAGGGTGGATCACCTG	21302
Qy	556	-----	556
Db	21301	AGGTCAAGATTCAAGATCAGCTTAGTCAACATGTTGAAACCCAGTCTCTACTAAATATA	21242
Qy	556	-----	556
Db	21241	CAAAAATTAGTGGGCACCGTGTGGGCACCTGTAATCCAGCTACTCGGCGGCTGAGG	21182
Qy	556	-----	556
Db	21181	CAGGAGAAATCGCTTGAAATCCGGGAGCGGAGGTTGCAGTGTAGTCAGATCACACCATGT	21122
Qy	556	-----	556
Db	21121	GCTCGGCTGGGAAACAAGACGCGAAACTTCGTCTAAAAAAGAAAGAAAGAAAT	21062
Qy	557	-----Gly-IleAsnValHisThrGluAspLeuLeuArgAsnProSerTyIleGlnGlu	574
Db	21061	CTACTGGAAATAAATGTACATCTCAAGACTTGCTCAGAAACCCATCTATATTCAAGAG	21002
Qy	575	AlaLysAlaLysGlyLeuValIlePheCysTrpGlyAspAspThrAsnAspProGluAsn	594
Db	21001	GCAAAAGCTAAGGAGCTAATCATATTTCTGCTGGGGTGATGAAGCAGTGCCTGAAAAG	20942
Qy	595	ArgArgLysLeuLysGluLeuGlyValAsnGlyLeuIleTyAspArgIleTyArgTrp	614
Db	20941	AGAAGAAATTAATGGAACCTTGAGTTAATGCTTATTTATCATAGGATATATGATTGG	20882
Qy	615	MetProGluGlnProAsnIlePheGlnVal-GluGlnLeuGluArgLysGlnGluLe	634
Db	20881	ATGCTCTGAACAACCA-----AGTGAGCACTGGAATGCTGAAGCAAGAATT	20834
Qy	634	uProGluLeuLysSerCysLeuCySProThrValSerArgPheValProSerSerLeuCy	654
Db	20833	GCCAGAGTGAAGGGCTGTTGTTGTTGCCACTGTGAGCCACTTTGTTCCCTCATCTTTGG	20774
Qy	654	sGlyGluSerAspIleHisValAspAlaAsnGlyIleAspAsnValGluAsn	671
Db	20773	TGAGGAGCTGTATCCCATGTGGATGCCAATGCAATGATATGAGGAAT	20722
RESULT	11		
LOCUS	HSN804380	6671 bp mRNA linear PRI 13-MAY-2003	
DEFINITION	Homo sapiens mRNA; cDNA DKFZp451O1715 (from clone DKFZp451O1715).		
ACCESSION	AL833069		
VERSION	AL833069.1 GI:21733660		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	1 (bases 1 to 6671)		
AUTHORS	Bloeker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764		
COMMENT	Neuberberg, GERMANY		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer		
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;		

sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp451O1715) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES
Location/Qualifiers
1..6671
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp451O1715"
/db_xref="taxon:9606"
/clone="DKFZp451O1715"
/tissue_type="human spinal cord"
/clone_lib="451 (synonym: hlcc1). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
6630..6635
polyA_signal
6645
polyA_site

ORIGIN

Alignment Scores:
Pred. No.: 9,34e-249 Length: 6671
Score: 2979.00 Matches: 569
Percent Similarity: 99.82% Conservative: 1
Best Local Similarity: 99.65% Mismatches: 1
Query Match: 84.22% Indels: 0
DB: 9 Gaps: 0

US-10-047-855-3 (1-672) x HSM804380 (1-6671)

Qy	102	LeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGluThr	121
Db	1713	CTAGAAAGCGAAATTTATTATTGACATGGAGCAATTTGGAAATCCCAATGGTGTGAAACT	1772
Qy	122	LeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTySerGlu	141
Db	1773	CTGGATTCTGGATGGCTGACATGTACACTGAAATAAGATTACGTTTGCATTATTCGAA	1832
Qy	142	LysProValSerIleThrLysLysLysLeuLysLysSerArgPheArgValLysLeu	161
Db	1833	AAACCTCTCTGTCATTAACCAAGAAAAATTAATAAAATCTAGATTAGGTGAAGCTG	1892
Qy	162	ThrLeuGluGlyLeuGluGluAspAspArgValSerProThrValLeuHisLys	181
Db	1893	ACACTAGAAGCCCTGGAGGAAGATGACGATGATAGGGTATCTCCCACTGTACTCCACAA	1952
Qy	182	MetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHisSer	201
Db	1953	ATGTCCAATAGCTTGGAGATATCCTTAATAAGCCACAATGAGTTCAAGTGCAGGCAITCA	2012
Qy	202	GlnProGluCysGlyTyTrpGlyLeuGlnProAspArgTrpThrGluTyTrSerIleGlnThr	221
Db	2013	CAGCCGGAGTGTGGTTATGGCTTGCAGCTGATCGTTGCACAGATCAGATACAGACG	2072
Qy	222	MetGluProAspAsnLeuGluIlePheAspPhePheGluGluAspLeuSerGluHis	241
Db	2073	ATGGNACCAGATTAACCTGGAATACTTTGATTTTTCGAAGAAGATCTCAGTGAGCAC	2132
Qy	242	ValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThr	261
Db	2133	GTAGTTTCAGGGTGATGCCCTTCTCGACATGTGGGTACAGCTTGTCTTATCATCCACC	2192
Qy	262	IleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsnSer	281
Db	2193	ATTGCTGAGAGTGGAAAGAGTGTCTTCTCTCCCATCATGAGCGAAATATCC	2252
Qy	282	ArgLysThrIleGlyLysValArgValAspTyIleIleIleLysProLeuProGlyTyTr	301
Db	2253	CGGAAAACATAGGCAAGTGGAGTTGACCATATATTAATTAAGCCATTACCAGGATAC	2312

```
QY 302 SerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgLysProLeuAspVal 321
Db 2313 AGTTGTGACATGAAATCTCTCATTTTCAAGTATTGGAAGCCCAAGAAATACCATTTGGATGTT 2372
QY 322 GlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAsn 341
Db 2373 GGCCATCGAGTGCAGGAACTCTACACAACTGCCAGCTGCTTAAGTTCAAGAAAT 2432
QY 342 ThrIleAlaSerLeuArgAsnAlaLysSerHisGlyAlaAlaPheValGluPheAspVal 361
Db 2433 ACTATTGCTCTTTTAAGAAATGCTGCTAGTCATGCTGAGCCCTTGTAGAAATTTGACGTA 2492
QY 362 HisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMet 381
Db 2493 CACCTTTCAAGAGACTTTGTGCCCGGTATATCATGATCTTACCTGTTGTTGACTATG 2552
QY 382 LysLysLysPheAspAlaAspProValGluLeuPheGluLysProValLysGluLeuThr 401
Db 2553 AAAAAGAAATTTGATGCTGATCCAGTTGATTTGAATTTCAAGTAAAGAAATTAACA 2612
QY 402 PheAspGlnLeuGlnLeuLysLysLeuThrHisValThrAlaLeuLysSerLysAspArg 421
Db 2613 TTTGACCAACTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTGAAATCTAAGGATCG 2672
QY 422 LysGluSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSerLeu 441
Db 2673 AAGAATCTGTGGTTCCAGGAGAAATTCCTTTTCAGAAATCAGCCATTTCTCTCT 2732
QY 442 LysMetValLeuGluSerLeuProGluAspValGlyPheAsnLysGluLysLysTrpIle 461
Db 2733 AGATGGTTTGAAGTCTTTGCCAGAGATGATAGGGTTTAAACATTAAGAAATAAATGGATC 2792
QY 462 CysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeu 481
Db 2793 TGCCAGCAAAAGGATGGATGGATGGATGTTATCAACATATTTTGACATGATCTG 2852
QY 482 PheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgLysValPhe 501
Db 2853 TTTTGTGATATATTTTAAAAACTGTTTTAGAAAATTTCTGGGAAGAGGAGAAATAGTGT 2912
QY 502 SerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIle 521
Db 2913 TCTTCAATTTGATGACATATTTGCACATGGTTTGAAGATCTACTGGGGATAAATGTCAT 2972
QY 522 LeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArg 541
Db 2973 CTATTTTAACTCAAGAAATCTGAGATTTATCTGTAACCTCATGACCTCAGATCTCG 3032
QY 542 ThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeuLeuGlyIleAsnValHis 561
Db 3033 ACAACCCCATTTGCAATGAGCTTTGCACAGTTTGAAGATCTACTGGGGATAAATGTCAT 3092
QY 562 ThrGluAspLeuLeuArgAsnProSerTyrIleGlnGluAlaLysLysGlyLeuVal 581
Db 3093 ACTGAAGACTGCTCAGAAACCCATCCTATTTCAAGAGCGCAAAAGCTAAGGACTAGTC 3152
QY 582 IlePheCysTrpGlyAspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeu 601
Db 3153 ATATTCTGCTGGGTGATGATACCATGATCTCTGAAAACAGAGGAATTTGAAGAACTT 3212
QY 602 GlyValAsnGlyLeuIleTyrAspArgLysTyrAspTrpMetProGluGlnProAsnIle 621
Db 3213 GGAGTTAATGGTCTAATTTTATGATAGATATATGATTGGATGCTCAACCAACCAATATA 3272
QY 622 PheGlnValGlnGlnLeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCysLeu 641
Db 3273 TTCCAAGTGGAGCAATTTGAACCCCTGGAAGCAGGAATTCACAGAGCTTAAGAGCTGTTG 3332
QY 642 CysProThrValSerArgPheValProSerSerLeuCysGlyGluSerAspIleHisVal 661
Db 3333 TGTCCACACTGTTAGCCGCTTGTTCCTCATCTTTGTGTGGGAGTCTGATATCCATGTG 3392
QY 662 AspAlaAsnGlyIleAspAsnValGluAsnAla 672
```

```
Db 3393 GATCCCAACGCGATTGATAACGTGGAGATGCT 3425
RESULT 12
AC150686/c
LOCUS
DEFINITION
AC150686
AC150686.1 GI:50872649
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS
Pan troglodytes (chimpanzee)
SOURCE
Pan troglodytes
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 160762)
AUTHORS
Wilson,R.K.
TITLE
The sequence of Pan troglodytes clone
JOURNAL
Unpublished
2 (bases 1 to 160762)
AUTHORS
Wilson,R.K.
TITLE
Direct Submission
JOURNAL
Submitted (31-JUL-2004) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: C.AB0549B03
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Consensus quality: 156619 bases at least Q40
Consensus quality: 157310 bases at least Q30
Consensus quality: 157914 bases at least Q20
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 8714: contig of 8714 bp in length
* 8715 8814: gap of unknown length
* 8815 35261: contig of 26447 bp in length
* 35262 35361: gap of unknown length
* 35362 62200: contig of 26839 bp in length
* 62201 62300: gap of unknown length
* 62301 114137: contig of 51837 bp in length
* 114138 114237: gap of unknown length
* 114238 158156: contig of 43919 bp in length
* 158157 158256: gap of unknown length
* 158257 159728: contig of 1372 bp in length
* 159729 160762: gap of unknown length
* 159729 160762: contig of 1034 bp in length.
FEATURES
Location/Qualifiers
1..160762
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="7"
/clone="CH251-549B3"
1..8714
/misc_feature
1..8714
/note="assembly_name:Contig16"
```

```
misc_feature      clone_end:SP6
                  vector_side:left"
8815..35261
/note="assembly_name:Contig17"
misc_feature      35362..62200
/note="assembly_name:Contig18"
misc_feature      62301..114137
/note="assembly_name:Contig19
clone_end:17"
vector_side:left"
114238..158156
/note="assembly_name:Contig20"
misc_feature      158257..159628
/note="assembly_name:Contig8"
misc_feature      159729..160762
/note="assembly_name:Contig9"

ORIGIN
Alignment Scores:
Pred. No.:      1,12e-246      Length:      160762
Score:          2976.50      Matches:      607
Percent Similarity: 78.00%      Conservative: 17
Best Local Similarity: 75.88%      Mismatches:   39
Query Match:     84.15%      Indels:       139
DB:              2          Gaps:           6

US-10-047-855-3 (1-672) x AC150686 (1-160762)
QY      1  MetThrProSerGlnValAlaPheGluIleArgGlyThrLeuLeuProGlyGluValPhe 20
DB      133550  ATGACACCTTCTCAGGTACCTTTGAATATAGAGGAATCTTTACCA--GAAGTTTTT 133494
QY      21  AlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAlaValAlaLeu 40
DB      133493  GCAATATGTGGAAGCTGTGATCTTTGGAAACTGCAATCTCTAAATGCTGTGGCTCTT 133434
QY      41  LeuProGluAsnAspThrGly-Glu----- 48
DB      133433  CTTCCAGAGATGAGACAGGTTGGAAGAGCTAAGATTGTATCAGAAATGAGGCAATGCCC 133374
QY      49  -----SerMetLeuTrpLysAlaThrIleValLeuSerArgG1 61
DB      133373  TAAACCTGTCTTACGACAGCCAGCATGTTATGGAAACCAACCATTTTACTCAGTAGAGG 133314
QY      61  yValSerValGlnTrpArgTrpPheLysGlyTyrPheLeuGluProLysThrIleGlyG1 81
DB      133313  AGTATCAGTCAGCATCTGCTACTTCAAAGGGTACTTTTTAGAACCAAGACTATCGGTGG 133254
QY      81  yProCysGlnValIleValHisLysTrpGluThrHisLeuGlnProArgSerIleThrPr 101
DB      133253  TCCATGTCAAGTGATAGTTTCAAGTGGGAGATTCTATACCAACCACCATCAATAACCCC 133194
QY      101  oLeuGluSerGluIleIleAspAspGlyGlnPheGlyIleHisAsnGlyValGluTh 121
DB      133193  GTCGAAGAGTGAAATGTTATTTGACCATGGCAATTTGGAATCCACAGTGGTGTGAAAC 133134
QY      121  rLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLeuHisTyrSerG1 141
DB      133133  TCTGGATTCTGCATGGCTGGCATGTGCACATGAATAGATTACGTTTGCATTATTCTGA 133074
QY      141  uLysProProValSerIleThrLysLysLeuLysLysSerArgPheArgValLysLe 161
DB      133073  AAAACCTCTGTGTCAATAACCAAGAAAAAATTTAAAAATCTAGATTAGGGTGAAGCT 133014
QY      161  uThrLeuGluGlyLeuGluAspAspAspArgValSerProThrValLeuHisIly 181
DB      133013  GACACTAGAGGCGCTGGGAGGAAGATGACATGATAGGGTATCTCCCATCTACTTCAAA 132954
QY      181  sMetSerAsnSerLeuGluIleSerLeuIleSerAspAsnGluPheLysCysArgHisSe 201
DB      132953  AATGTCCAAATAGCATGGAGATATCTTAAATATGTGACAAATGATGATTCAGTCAGGCATTC 132894
QY      201  rGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGlnTh 221
DB      132893  ACAGCTGGAGTGTGGTTATGCTTTACAGTCTGATCGTTGGACAGAGTACAGCATACAGAC 132834
QY      221  rMetGluProAspAsnLeuGluLeuPheAspPheGluGluAspLeuSerGluHi 241
DB      132833  AATGGGACCATTAACCTGGAACTAAATCTTTGATTTTGTGAAGAAATCTCAGTGACCA 132774
QY      241  sValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerTh 261
DB      132773  CATAGTTCAAGGTGATGCCCTCCCTGGACGTGGGTACAGCTGTCTCTATTATTCAC 132714
QY      261  rIleAlaGluSerGlyLysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsnSe 281
DB      132713  CATTTGCTGAGAGTGGAAAGAGTGTGGAAATTTCTACTCTCCCTACATGACGAGAAATTC 132654
QY      281  rArgLysThrIleGlyLysValArgValAspTyrIleIleIleLysProLeuProGlyTy 301
DB      132653  CAGGAAAAACAATAGCAAAAGTGAGTGCATATATAATATTATTAAGCCATTGCCAGGATA 132594
QY      301  rSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIleProLeuAspVa 321
DB      132593  CAGTTGTGACATGAAATCTTCAATTTCCAG-TATTGGGAAGCCAAAGATACCACCTGGATGT 132535
QY      321  lGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAs 341
DB      132534  TGGCCATCGAGGTGTGGGAAC---ACTACAACTGCCCGCTGGCTGCTAAA-----GAAAA 132484
QY      341  nThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAspVa 361
DB      132483  TACTATTGCTCTCTTTAAGAAATGCTGTAGTCACAGTCAGCCTTTGTAGAATTTGATGT 132424
QY      361  lHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCysCysLeuThrMe 381
DB      132423  ACACCTTTCAAAGAGCTTTGTGCCCATGATATCATGATCTTACCTGTGTGTGACTAT 132364
QY      381  tLysLysLysPheAspAlaAspProValGluLeuPheGluIleProVal-LysGluLeuT 401
DB      132363  GAAAAAGAAATTTGATGCTGATCCACTTGAATTTTGGAAATTCAGTAAATAAAATTTAA 132304
QY      401  hrPheAspGlnLeuGlnLeuLysLeuThrHisValThrAlaLeuLysSerLysAspA 421
DB      132303  CATTTGACCATCTCCAGTTGTTAAAGCTCACTCATGTGACTGCACTGAAATCTAAGGATT 132244
QY      421  rGlyLysLeuSerValValGlnGluAsnSerPheSerGluAsnGlnProPheProSerL 441
DB      132243  GGAACCAATCTGTGTCTCCAGGAGGAAATTCCTTTTCAGAAAAATCAGGACATTTCTCTC 132184
QY      441  euLysMetValLeuGluSerLeuProGluAspValGlyPheAsnIleGluIleLysTrpI 461
DB      132183  GTAAGATGGTTTTAGAGTGTGGCCAGAAAGATGAGGGTTTTAACATTGAAATAAAATGGA 132124
QY      461  leCysGlnGlnArgAspGlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnL 481
DB      132123  TCTGCCCAAGGAGTGAATGTGGATGTGAATCTTATCAACATATTTTGGCATGAATC 132064
QY      481  euPheLeuAspIleIleLeuLysThrValLeuGluAsnSerGlyLysArgArgIleValP 501
DB      132063  TGTGTTTTGGATATAACTTTTAAAAAAGCTGTTTTAGAAAAATCTCGGGAAGAG-AGAAATAGTGT 132005
QY      501  heSerSerPheAspAlaAspIleCysThrMetValArgGlnLysGlnAsnLysTyrProI 521
DB      132004  TTTCTCATTTGATGCGATATTTGCACAAATGGTTTGGCAAGAGCAGAAACAAATATCCCA 131945
QY      521  leLeuPheLeuThrGlnGlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerA 541
DB      131944  TATTATTTTAACTGAAGGAAAAATCTGAGATTTATCTGAACTCATGACCTCAGATCTC 131885
QY      541  rgThrThrProIleAlaMetSerPheAlaGlnPheGluAsnLeu----- 556
DB      131884  CGACAAACCCCATTTGCAATGAGCTTTGCACAGTTTGAATAATCTACTAGGATAGGCTGGG 131825
QY      556  ----- 556
```

```
Db 131824 TCGCGGGCTCGCGCTGTAAACCCAGCACTTTGGGAGCGCGAGGAGGTGGATCACCTG 131765
QY 556 -----
Db 131764 AGGTCAGAGTTCAAGATCAGCTTAGTCAACATGGTGAACCCAGTCTCTACTAAATATA 131705
QY 556 -----
Db 131704 CAAAAATTAGCGGCGCACCGTGGTGGCACCTGTAATCCAGCTACTCGGCAGGCTGAGG 131645
QY 556 -----
Db 131644 CAGGAGATCGCTTGAATCCGGAGCGGAGGTTCAGTCAGTCGAGATCACGCCATTGT 131585
QY 556 -----
Db 131584 GCTCCGGCTGGGAACAAGAGCGAAACTTCGCTTAAAAAAGAAAAAGAAAAAT 131525
QY 557 -----Gly-IleAsnValHisThrGluAspLeuLeuArgAsnProSerTyrIleGlnGlu 574
Db 131524 CTACTGGAGATAAATCCACATCTGAAGACTTGCTCAGAAACCCATCTACATTCAGAG 131465
QY 575 AlaIleAlaIleGlyLeuValIlePheCysTyrGlyAspAspThrAsnAspProGluAsn 594
Db 131464 GCAAAAGCTAAGGACTAATCATATTCTGCTGGGGTGATGACAGCATGATCCCGAAAG 131405
QY 595 ArgArgLeuLeuGlyLeuGlyValAsnGlyIleIleTyrAspArgIleTyrAspTyr 614
Db 131404 AGAAGAAATTTGAAGAACTTGGAGTTAATGGTCTATTATGATAGGATATATGATTGG 131345
QY 615 MetProGluGlnProAsnIlePheGlnVal -GluGlnLeuGluArgLeuLeuGlyGlnGlu 634
Db 131344 ATGCTTGAACCAACA-----AGTGGAGCAACTGGAATGCTGAAGCAAGAAAT 131297
QY 634 uProGluLeuLeuSerCysLeuCysProThrValSerArgPheValProSerSerLeuCy 654
Db 131296 GCCAGAGTAAAGGCTGTTTGTGCTCCACTGTGACCACTTTGTTCCCTCATCTTTGTG 131237
QY 654 sGlyGluSerAspIleHisValAspAlaAsnGlyIleAspAsnValGluAsn 671
Db 131236 TGAGGAGCTGATACCATGTGGATGCCAATGGCAATTGATATGTGAGAAAT 131185

RESULT 13
CQ728047
LOCUS 1731 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 13981 from Patent WO02068579.
ACCESSION CQ728047
VERSION CQ728047.1 GI:42295769
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 13981 06-SEP-2002;
FE Corporation (NY) (US)
FEATURES
source
1. .1731
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 5,65e-246 Length: 1731
Score: 2938.00 Matches: 575
Percent Similarity: 87.65% Conservative: 0
Best Local Similarity: 87.65% Mismatches: 0
Query Match: 83.06% Indels: 81
```

```
DB: 6 Gaps: 3
US-10-047-855-3 (1-672) x CQ728047 (1-1731)
QY 18 GluValPheAlaIleCysGlySerCysAspAlaLeuGlyAsnTrpAsnProGlnAsnAla 37
Db 3 GAAGATTTTTCGATATGTGAAGCTGTGATCTTTGGAAACTGGAAATCTCTCAAAATGCT 62
QY 38 ValAlaLeuLeuProGluAsnAspThrGlyGluSerMetLeuTrpLysAlaThrIleVal 57
Db 63 GTGGCTCTTCTCCAGAGAATGACACAGGTGAAGCATGCTATGGAAGCAACCATTTGTA 122
QY 58 LeuSerArgGlyValSerValGlnTyrArgTyrPhe -LysGlyTyrPheLeuGluProLy 77
Db 123 CTCAGTAGAGGAGTATCAGTTTCACTATCGCTACTTCAAG----- 162
QY 77 sThrIleGlyGlyProCysGlnValIleValHisTyrTrpGluThrHisLeuGlnProAr 97
Db 162 ----- 162
QY 97 gSerIleThrProLeuGluSerGluIleIleIleAspAspGlyGlnPheGlyIleHisAs 117
Db 163 -----GAAAGCGAATTTATTGACGATGGACAATTTTGAATCCACAA 206
QY 117 nGlyValGluThrLeuAspSerGlyTrpLeuThrCysGlnThrGluIleArgLeuArgLe 137
Db 207 TGGTGTGAACCTCTGGATTTCTGGATGCTGACATGTCAGACTGAAATAAGATTACGTTT 266
QY 137 uHisTyrSerGluLeuProValSerIleThrIleThrIleLeuLeuLeuLeuLeuLeu 157
Db 267 GCATTATTCTGAAAAACCTCTGTGTCAATAACCAAGAAAAAATTTAAAAAATCTAGATT 326
QY 157 eArgValLeuLeuThrLeuGluGlyLeuGluAspAspAspAspArgValSerProTh 177
Db 327 TAGGGTGAAGCTGACACACTAGAGGCTCGAGGAGATGACGATGATAGGGTATCTCCAC 386
QY 177 rValLeuHisLysMetSerAsnSerLeuGluLeuSerLeuIleSerAspAsnGluPheLy 197
Db 387 TGTACTCCACAAATATGCTCAATAGCTTGGAGATATCTTTAATAAGCGACAATGAGTTCAA 446
QY 197 sCysArgHisSerGlnProGluCysGlyTyrGlyLeuGlnProAspArgTrpThrGluTy 217
Db 447 GTGAGGCAATTCACAGCCGAGTGTGTTATGGCTTGCAGCCCTGATCGTTGGACAGAT 506
QY 217 rSerIleGlnThrMetGluProAspAsnLeuGluLeuIlePheAspPhePheGluGluAs 237
Db 507 CAGCATACAGACGATGGAACACAGATACTGGNACTAATCTTTGATTTTTCGAAGAGA 566
QY 237 pLeuSerGluHisValValGlnGlyAspAlaLeuProGlyHisValGlyThrAlaCysLe 257
Db 567 TCTCAGTGAGCAGCTAGTTTCAGGGTGTGTCCTTCCTGGACATGTGGGTACAGCTTGTCT 626
QY 257 uLeuSerSerThrIleAlaGluSerGlyLysSerAlaGlyIleLeuThrIleuProIleMe 277
Db 627 CTTATCATCCACCATTTCTGAGGTGGAAGAGTGTCTGAAATCTTTACTCTTCCCATCAT 686
QY 277 tSerArgAsnSerArgLeuThrIleGlyLysValArgValAspTyrIleIleIleLysPr 297
Db 687 GAGCAGAAATTCGGGAAACCAATAGGCAAGTAGAGTTGACTATATATATTATTAAAGCC 746
QY 297 oLeuProGlyTyrSerCysAspMetLysSerSerPheSerLysTyrTrpLysProArgIl 317
Db 747 ATTACCAGGATACAGTTGTGACATGAAATCTTTCATTTTCCAAGTATTGGAAGCAAGAA 806
QY 317 eProLeuAspValGlyHisArgGlyAlaGlyAsnSerThrThrAlaGlnLeuAlaLy 337
Db 807 ACCATTGGATTTGGCCATCGAGGTGAGGAACTCTCAACAACTGCCCGAGCTGCTTAA 866
QY 337 sValGlnGluAsnThrIleAlaSerLeuArgAsnAlaAlaSerHisGlyAlaAlaPheVa 357
Db 867 AGTTCAAGAAATAATCTATTGCTTTCTTAAGAAATGCTGCTAGTCATGTTGGACGCTTTGT 926
QY 357 lGluPheAspValHisLeuSerLysAspPheValProValValTyrHisAspLeuThrCy 377
```



```

Qy 618 GlnProAsnIlePheGlnValGluGlnLeuGluArgLeuLysGlnGluLeuProGluLeu 637
Db 1793 CAGCAATAATATTCAGAGTGAGCGATTGAGCGCTGAGCAGAGATTGCCAGAGCTT 1852
Qy 638 LysSerCysLeuCysProThrValSerArgPheValProSerSerLeuCysGlyGluSer 657
Db 1853 AAGAACTGTTGTGCTCCACTGTAGCCACTTCATCCCTTCTTCTTCTGTGTGGAGCT 1912
Qy 658 AsnIleHisValAspAlaAsnGlyIleAspAsnValGluAsnAla 672
Db 1913 GATATCCACGTGGATGCCAAGCGCATTCATAGTGTGGAGAGCT 1957

RESULT 15
AX127525 AX127525 2738 bp DNA linear PAT 15-MAY-2001
LOCUS Sequence 1 from Patent WO0131007.
DEFINITION AX127525
ACCESSION AX127525
VERSION AX127525.1 GI:14134230
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chiang, L.W.
TITLE Nucleic acid molecules derived from rat brain and programmed cell
death models
JOURNAL Patent: WO 0131007-A 1 03-MAY-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..2738
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 4,35E-212 Length: 2738
Score: 2552.00 Matches: 486
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.15% Indels: 0
DB: 6 Gaps: 0

US-10-047-855-3 (1-672) x AX127525 (1-2738)
Qy 187 GluIleSerIleSerAspAsnGluPheLysCysArgHisSerGlnProGluCysGly 206
Db 18 GAGATATCCTTAATAGCGACAATGAGTTCAAGTGCAGGCATTCACAGCGAGTGTGT 77
Qy 207 TyrGlyLeuGlnProAspArgTrpThrGluTyrSerIleGlnThrMetGluProAspAsn 226
Db 78 TATGGCTTCGAGCTGATCGTTGGACAGAGTACAGCATACAGCATGACAGATAC 137
Qy 227 LeuGluLeuIlePheAspPheGluGluAspLeuSerGluHisValGlnGlyAsp 246
Db 138 CTGGAACATACTCTTGATTTTCGAGAAGATCTCAGTGAGCACGCTAGTTCAGGGTGAT 197
Qy 247 AlaLeuProGlyHisValGlyThrAlaCysLeuLeuSerSerThrIleAlaGluSerGly 266
Db 198 GCCCTTCCTGGACATGTGGGTACAGCTTGTTCTTATCATCCACCATTGCTGAGAGTGA 257
Qy 267 LysSerAlaGlyIleLeuThrLeuProIleMetSerArgAsnSerArgLysThrIleGly 286
Db 258 AAGAGTGTGGAAATCTTACTCTTCCCATCATGAGCAGAAATCCCGGAAACATATGCC 317
Qy 287 LysValArgValAspTrpIleIleLysProLeuProGlyTyrSerCysAspMetLys 306
Db 318 AAGTGTGAGTGTGACTATATAATTAAGCCATTACCCAGGATACAGTGTGTGATGAAA 377
Qy 307 SerSerPheSerLysTyrTrpLysProArgIleProLeuAspValGlyHisArgGlyAla 326

```

```

Db 378 TCCTCATTTTCCAAGTATTGGAAAGCCAGAAATACCATTTGGATTTGGCCATCGAGGTGCA 437
Qy 327 GlyAsnSerThrThrAlaGlnLeuAlaLysValGlnGluAsnThrIleAlaSerLeu 346
Db 438 GGAACACTCTACACACACTGCCAGCTGCTGCTCAAGAAATACTATTGCTTCTTTA 497
Qy 347 ArgAsnAlaAlaSerHisGlyAlaAlaPheValGluPheAspValHisLeuSerLysAsp 366
Db 498 AGAAATGCTGCTAGTCACTGCTGAGCCCTTTGTAGAATTTGACGTACACCTTTCAAGAGAC 557
Qy 367 PheValProValValTyrHisAspLeuThrCysCysLeuThrMetLysLysLysPheAsp 386
Db 558 TTTGTGCCCGTGGTATATCATGATCTTACCTTGTTGTTGACTATCAAAAGAAATTTGAT 617
Qy 387 AlaAspProValGluLeuPheGluIleProValLysGluLeuThrPheAspGlnLeuGln 406
Db 618 GCTGATCCAGTTGAAATTTTGAATTTCCAGTAAAGAAATTAACATTTGACCAACTCCAG 677
Qy 407 LeuLeuLysLeuThrHisValThrAlaLeuLysSerLysAspArgLysGluSerValVal 426
Db 678 TTGTTAAAGCTCACTCATGTGACTGCACTGAAATCTAAGGATCGGAAAGAAATCTGTGTT 737
Qy 427 GlnGluGluAsnSerPheSerGluAsnGlnProPheProSerLeuLysMetValLeuGlu 446
Db 738 CAGGAGGAAATTCCTTTTCAGAAATTCAGCCATTTCTTCTCTTAAGATGGTTTAGAG 797
Qy 447 SerLeuProGluAspValGlyPheAsnIleGluIleLysTrpIleCysGlnGlnArgAsp 466
Db 798 TCTTTGCCAGAGATGTAGGGTTTAAACATTTGAAATTAATGATCTGCCCAAGGAT 857
Qy 467 GlyMetTrpAspGlyAsnLeuSerThrTyrPheAspMetAsnLeuPheLeuAspIleIle 486
Db 858 GGAATGGGATGTTAACTATCAACATATTTTGCATGATCTGTTTTCGATATAATT 917
Qy 487 LeuLysThrValLeuGluAsnSerGlyLysArgArgIleValPheSerSerPheAspAla 506
Db 918 TTAATAAACTGTTTAAAGAAATTCCTGGAGAGGAGAAATAGTGTTCCTTCATTGATGCA 977
Qy 507 AspIleCysThrMetValArgGlnLysGlnAsnLysTyrProIleLeuPheLeuThrGln 526
Db 978 GATATTTGCACAAATGGTTTCGCAAAAGCAGAAATAATCCGATACTATTTTACTCAA 1037
Qy 527 GlyLysSerGluIleTyrProGluLeuMetAspLeuArgSerArgThrProIleAla 546
Db 1038 GGAATACTGAGATTTATCTGAACTCATGACCTCAGATCTCGACAAACCCCATTCGA 1097
Qy 547 MetSerPheAlaGlnPheGluAsnLeuGlyIleAsnValHisThrGluAspLeuLeu 566
Db 1098 ATGAGCTTTGCACAGTTTGAATACTACTGGGATAAATGTACATACTGAAAGACTTGCTC 1157
Qy 567 ArgAsnProSerTyrIleGlnGluAlaLysAlaLysGlyLeuValIlePheCysTrpGly 586
Db 1158 AGAAACCCATCCTATATTCAAGAGGCAAAAGCTAAGGGACTAGTCATATTTCTGCTGGGT 1217
Qy 587 AspAspThrAsnAspProGluAsnArgArgLysLeuLysGluLeuGlyValAsnGlyLeu 606
Db 1218 GATGATACCAATGATCTCGAAACAGAAAGAAATTCAGAGAACTTGGAGTTAATGCTCTA 1277
Qy 607 IleTyrAspArgIleTyrAspTrpMetProGluGlnProAsnIlePheGlnValGluGln 626
Db 1278 ATTTATGATAGATATATGATTTGGATGCTGAAACAACAAATATATTCCAAGTGGAGCAA 1337
Qy 627 LeuGluArgLeuLysGlnGluLeuProGluLeuLysSerCysLeuCysProThrValSer 646
Db 1338 TTGGAACGCTCAGACAGCAATTCAGAGCTTAAGAGCTGTTTGTGCTCCACTGTATGC 1397
Qy 647 ArgPheValProSerSerLeuCysGlyGluSerAspIleHisValAspAlaAsnGlyIle 666
Db 1398 CCGCTTGTTCCTCATCTTTGTGTGGGAGTCTGATATCAATTCATGTGATGCCAACCGCAT 1457
Qy 667 AspAsnValGluAsnAla 672
Db 1458 GATACGTGGAGAAATGCT 1475

```

Search completed: July 3, 2005, 09:24:08
Job time : 8255 secs

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)